Comparative Genomics to Identify Genetic Variabilities Associated with *Salmonella* Typhimurium DT104 Epidemicity

Mark P. Saunders

Royal Holloway University London
School of Biological Sciences
MPhil Thesis

Veterinary Laboratories Agency
Supervisor – Prof. M. J. Woodward.

Royal Holloway
Supervisor – Prof. P. Bramley.

16-10-2012

Declaration of Authorship

I ……..... hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Signed: _____

Date: _____

Abstract

Non-typhoidal *Salmonella* epidemics have a huge impact on animal and human health. During the 1980's a new epidemic strain emerged, *S.* Typhimurium DT104, first in the UK but later spreading through Europe and North America. This phage type displayed resistance to a panel of antibiotics and was isolated from cattle, pigs and poultry as well as from human clinical infections. By the 2000's, infections by this strain were subsiding, while other epidemic strains of *Salmonella* were emerging.

This study used Comparative Genomic Hybridisation (CGH) arrays to screen DT104 strains from before, during and after the epidemic period, to interrogate the genetic changes that may have contributed to the strain epidemicity. Two novel techniques, optical mapping and the Biolog phenotypic microarray, were used to confirm or refute the CGH results. The study found that the major difference between pre-epidemic and epidemic strains was the acquisition of the antibiotic resistance carrying *Salmonella* Genomic Island 1 (SGI1). This region showed increased variation towards the latter period. In addition, genes involved in the theorised production of a 'pertussis like' toxin were found to be present in human epidemic strains, while they varied or were missing from strains of animal origin and human pre-epidemic sensitive strains. Other differences were found in the presence of genes for allantoin and glyoxylate utilisation for strains that may represent an ancestral clone of DT104. The CGH results revealed variation and this variation was confirmed by phenotypic assay and confirmatory techniques.

Overall the picture of DT104 epidemicity is complex, but this study has identified areas of potential importance and has suggested avenues of research, that may lead to a deeper understanding of how epidemics develop.

Acknowledgments

I would first like to thank my two supervisors at the Veterinary Laboratories Agency (VLA), Dr Muna Anjum who was my first supervisor and Prof. Martin Woodward who supervised the latter part. Between them they have given me the help and support that made this thesis possible. I'd also like to thank my supervisor at Royal Holloway University London, Prof. Peter Bramley, who has helped facilitate the study and has always been available to aid with the administrative hoops and hurdles. I also like to express my gratitude to the VLA, for the funding of this project and the Wellcome Trust Sanger Institute for the preparing and providing the CGH microarrays.

The work would not have been possible without the help and support of colleagues. They include: Dr. Manal AbuOun, Mary Bagnall, Emma Kennedy, Dr. Roberto La Ragione and Dr. Guanghui Wu. I'd especially like to thank Manal for her help with the Biolog Phenotypic Microarray technique. I also thank Dr Zhensheng Pan who taught me the CGH microarray technique, and as such was indispensable.

But most of all I want to thank my family for their love and support, especially Mum and Dad, who put up with a student moving back in to write up, and provided help and guidance.

Mostly however I want to thank my wife Ellin, without who I would have never started this journey or survived its passage. Your love and support has been my rock, Jag älskar dig älskling.

Contents

Abbreviations

ATR	Acid tolerance response
BSAC	British Society for Antimicrobial Chemotherapy
C.F.U.	Colony-forming unit
CGH	Comparative genomic hybridisation
DT	Definitive type (alternative phage type moniker)
HPA	Health Protection Agency
LB	Luria-bertani (broth or agar)
LPS	Lipopolysaccharide
MDR	Multidrug resistance
MIC	Minimum inhibitory concentration
MLST	Multi-locus sequence typing
OD	Optical density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PFGE	Pulsed field gel electrophoresis
PMN	Polymorphonuclear neutrophillic leukocytes
PT	Phage type
ROI	Reactive oxidative intermediates
SGI1	*Salmonella* genomic island 1
SNPs	Single-nucleotide polymorphisms
SPI	*Salmonella* pathogenicity island
T1SS	Type one secretion system
TTSS	Type three secretion system
VLA	Veterinary Laboratories Agency

1 Introduction

This study considers some of the factors that underlie the emergence of epidemic *Salmonella enterica* strains, the mechanisms by which they become prevalent, the processes of their evolution and the implications of these processes on genetic characteristics and resultant phenotype with regard to developing potential management and control measures. The importance of discovering the reasons behind epidemics of *Salmonella* is underlined by the economic cost on both health services and to the farming industry. In 2007 the cost of human *Salmonella* infections was calculated at over $2.5 billion (ERS, 2008).

Classic epidemiology tells us that in broad terms there are three major components to consider, namely the organism, the host with which it interacts and the environment through which the organism passes. No one study can realistically cover all these aspects and so the focus here is upon gaining an understanding of the organism selected for study, *Salmonella enterica* subs. *enterica* ser. Typhimurium definitive type (DT) 104 that emerged as a significant epidemic type during the 1980's and throughout the 1990's – 2000's.

This chapter will introduce the fundamentals of *Salmonella* that will impact on this study as well as outline the aims of the project.

1.1 Overview of *Salmonella*

Salmonella is an important zoonotic pathogen linked with serious animal and human disease. It is one of the most common causes of food-borne infections in man (Plym Forshell, 2006). Typhoidal salmonellosis or typhoid fever was estimated to have caused 22 million illnesses and over 200,000 deaths in the year 2000. Paratyphoid fever caused 5 million cases for the same time period (Crump, 2004). Non-typhoidal salmonellosis which includes that caused by *S.* Typhimurium also has a high burden with an estimated 1.4 million cases in the USA

annually with 15,000 hospitalisations and 400 deaths (Mead, 1999; Voetsch, 2004). There were also over 17,500 cases reported to laboratories in the UK in 1999 (WHO, 2003).

The genus *Salmonella* is a member of the family *Enterobacteriaceae* and consists of a collection of closely related Gram-negative, rod shaped, non-spore forming bacterial species. Traditional methods of typing have relied on antigenic reaction to highly variable surface antigens expressed by the bacterium. These include the O-antigen of surface lipopolysaccharide (LPS) and the H-antigen of surface expressed flagella. Most *Salmonella* contain two flagellin genes which code for H1 and H2 variants of the H antigen, and serotyping has to include phase conversion to identify both types (Popoff, 2001).

Studies into nucleotide sequence relationships have shown that, typically, *Salmonella* are 85–100% related (Crosa, 1973). In addition multilocus enzyme electrophoresis (MLEE) studies have shown that *S. bongori* has a higher level of difference, with only a 43% similarity to other *Salmonella* strains (Reeves, 1989). The MLEE technique compares the electrophoretic mobility of core enzymes, and has a lower resolution than molecular techniques such as MLST. These approaches have resulted in the separation of *Salmonella* into two species *S. enterica* and *S. bongori*. *S. enterica* is further sub-divided into six subspecies, *enterica*, *salamae*, *arizonae*, *diarizonae*, *indica* and *houtenae*. They are classified respectively into subspecies I, II, IIIa, IIIb, IV and VI. *S. bongori* was previously designated as subspecies V. The White-Kauffmann-Le Minor scheme divides the subspecies further into 2541 serovars via their antigenic profile (Popoff, 2002). Over half of these are contained in *S. enterica* subspecies *enterica*, whereas the species *S. bongori* accounts for only 21 of the serovars (Popoff, 2001). MLEE has suggested an additional subspecies, VII, related to subsp. IV (Boyd et al., 2001) although this group is not distinguishable by unique biochemical properties. The phylogenetic relationship between the species and subspecies has been

resolved by analysing the nucleotide information from 4 houskeeping genes, *gapA*, *phoP*, *mdh* and *recA* (McQuiston, 2008).

Other methods of typing have been developed to further differentiate within serotypes, of which one is bacteriophage or 'phage' typing. The first scheme used on an international scale took advantage of adaptations of the Vi-phage II. This phage was specific to the Vi capsular antigen of *Salmonella enterica* subsp. *enterica* Typhi (Felix, 1934). Most phage typing schemes use a panel of serologically distinct phages isolated from a myriad of sources. Such schemes have been put into place for *Salmonella enterica* serovars *S.* Paratyphi, *S.* Hadar, *S.* Virchow, *S.* Enteritidis and *S.* Typhimurium. The characterisation of the phages for the *S.* Enteritidis phage typing scheme showed that the 16 phages used come from three morphotypes, *Podoviridae*, *Siphoviridae* and *Myoviridae* (De Lappe, 2009). The *S.* Typhimurium scheme started by Felix and Callow (Felix, 1943) and then expanded by Anderson (Anderson, 1977) has distinguished over 300 phage types for *S.* Typhimurium. This subdivision into phage types allowed the quicker differentiation of emerging strains and allowed their impact and spread to be measured.

The vast majority of human salmonellosis cases are caused by subspecies 1, 70% by just 20 serovars (CDC, 2008). The low proportion of 'problem' serovars is likely the product of the lack of contact with other types and also their host specificity. Multilocus sequencing typing (MLST) has been used to show that bovine and human associated *Salmonella* isolates represent distinct and overlapping population (Alcaine et al., 2006). Of the 10 serovars most commonly isolated from human infections, 6 are found in the top 10 serovars isolated from both swine and poultry (Foley, 2008).

1.1.1 *Salmonella* nomenclature

The nomenclature of *Salmonella* is complex as it has developed over time in response to diagnostic issues. For this report the scheme used by the Center for Disease Control will be followed (Brenner, 2000). Therefore the two species shall be cited as *S. bongori* or *S. enterica* and individual serovars shall be given either the full title of; *Salmonella enterica* subsp. *enterica* serovar Typhimurium or the shortened version of *S.* Typhimurium.

1.2 Host specificity

Different serovars of *Salmonella* display different levels of host specificity. The more host adapted strains, such as *S.* Gallinarum in poultry or *S.* Dublin in cattle (59% of cattle isolations in 2006 were *S.* Dublin (VLA, 2007)), are generally less pathogenic to humans, although there is some evidence that such infections can lead to more severe clinical symptoms. Between 1981 and 1990, 25% of human infections with *S.* Dublin developed septicaemia compared with 1-2% with the more common *S.* Typhimurium and *S.* Enteritidis infections (Threlfall, 1992). The two latter serotypes are the two most commonly isolated *Salmonella* isolates in humans. In the UK in 2000, *S.* Enteritidis accounted for 57% of *Salmonella* infections while *S.* Typhimurium accounted for 18% with the next highest contribution at 2% (WHO, 2003). By 2006 these two serovars still accounted for 66% of human isolations (VLA, 2007). During the 1980s there was a considerable increase in the number of human food-borne illness caused by *S.* Enteritidis; of 35 countries, 24 (69%) recorded an increase in *S.* Enteritidis isolates (Rodrigue, 1990).

The prevalence of serotypes in outbreaks also seems to vary with region. In the US in 2006 the top serovars in humans, responsible for over half of infections, were *S.* Typhimurium, *S.* Enteritidis and *S.* Newport. The serovars *S.* Typhimurium, *S.* Newport, *S.*

Agona and *S.* Orion were responsible for 38.3% of total isolations from farm animals showing clinical symptoms, while from non-clinical farm isolation (herd and flock monitoring, feed testing, environmental sampling etc) *S.* Heidelberg and *S.* Kentucky were the most commonly found (CDC, 2008). This is in contrast with the situation in Australia, where the top three human serovars for the same period were *S.* Typhimurium, *S.* Vichow and *S.* Saintpaul and from farm isolations *S.* Typhimurium, *S.* Dublin and *S.* Bovismorbificans (Powling, 2005). Variation within serovars is also apparent, in 2006 *S.* Newport accounted for 9% of UK cattle isolations and 1.7% of human isolations. However this serovar is of greater concern in the US where from 1996 to 2001 the human prevalence increased from 5% to 10% (Gupta, 2003). Normally isolated from cattle, Newport has emerged in the US as an issue affecting human health as it carries plasmids that confer antibiotic resistances and is developing resistances to later generation cephalosporins. In the UK, *S.* Newport does not demonstrate antibiotic resistance and has a variant pulsed field gel electrophoresis (PFGE) profile.

1.2.1 *Salmonella enterica* subsp. *enterica* serotype Typhimurium

S. Typhimurium (with the antigenic formulae 1,4,[5],12;i;1,2) is a zoonotic serotype commonly isolated from cattle although it is considered highly promiscuous as it is associated with disease or infection of a number of vertebrate species. It is associated with pyrexia, diarrhoea and in dairy cattle a drop in milk production (Sharp and Rawson, 1992). Over time certain phage types of *S.* Typhimurium have become predominant. During the 1960's multiple resistant phage type (PT) 29 became widespread in calves (Anderson, 1968). By 1977 PT204 and PT193 were most frequently isolated, with the variant definitive type

(DT) 204c [1] being the most common (Wray et al., 1998; Wray, 1987). Then in the late 80's and early 90's *Salmonella* Typhimurium DT104 emerged as a new epidemic strain.

1.3 Epidemiology

1.3.1 Initial colonisation

Salmonella Typhimurium causes gastroenteritis in humans and farmed animal species including cattle and pigs, whereas in mice it has been shown to cause a systemic disease similar to typhoid fever, and is therefore a surrogate model for the study of systemic typhoid. The bacteria are ingested via contaminated food or water. In human volunteers the minimum infectious dose is approximately 10^6, however single food outbreaks suggest that lower doses can cause disease. Bacteria that survive the low pH of the stomach locate to the lumen of the small intestine. In poultry, infected food reaches the crop, and subsequently the highest viable counts are found in the cecum, cloaca and ileum (Turner et al., 1998).

The interaction of *Salmonella* with the host environment and the subsequent triggering of acute inflammatory reaction causes ileal secretion and diarrhoea (Giannella, 1979). Interaction with protective cells such as polymorphonuclear leuocytes (PMN's) and macrophages, and the production of prostaglandin, and subsequently adenylate cyclase leads to the fluid secretion. *S.* Enteritidis and *S.* Typhimurium have also been shown to disrupt gastro-intestinal epithelial tight junctions, which then may trigger ion efflux and diarrhoea. The structural damage caused by *S.* Typhimurium may also lead to the release of damaging oxidative compounds

[1] PT (phage type) and DT (definitive type) arise from phage typing naming schemes from different countries and should be considered equivalent.

In the murine model and in immunocompromised adults non-typhoidal *Salmonella* can also cause systemic infection. Persistence in the reticuloendothelial system with spread to the bone marrow, liver and spleen can then lead to bacteraemia (Gordon, 2008).

1.3.2 Invasion

In mice *Salmonella* principally invade the murine M cells, specialised epithelial cells which line the Peyer's Patches of the intestine (Jones et al., 1994). The contact of *Salmonella* with the host cell membrane leads to changes in the brush border, with the microvilli becoming denatured. Membrane 'ruffles' begin to form and the bacteria are rapidly internalized inside a membrane bound vesicle (Jones et al., 1993). Distinct from other bacteria *Salmonella* remain inside the vesicle throughout the period of internalization. The environment within the vesicle has been shown to contain lysine and oxygen and low concentrations of Mg^{2+} and Fe^{2+} and has a mildly acidic pH (Garcia-del Portillo, 1992). Internalised bacteria later exit through the basal surface into the reticuloendothelial system where they are readily engulfed by macrophages.

Variations in the ability of different serovars to invade murine epithelia may be indicative of host specificity. *S.* Gallinarum is unable to invade murine M cells, while *S.* Typhimurium shows a greater ability to invade and persist then *S.* Typhi (Pascopella, 1995). The two avian adapted serovars *S.* Gallinarum and *S.* Pullorum can demonstrate increased invasiveness and the ability to cause membrane ruffles with mammalian cells when they are made to express *S.* Typhimurium type 1 fimbriae (Wilson et al., 2000).

1.3.3 Bacterial survival and general host defences

Salmonella have to avoid many adverse conditions to survive within a host. Firstly the bacteria have to overcome innate defences. *Salmonella* are sensitive to deviations from a neutral pH, and thus the conditions present in the gut are generally detrimental to its survival. It has developed mechanisms to counter-act the environment. Low pH has been shown to trigger an acid tolerance response (ATR) in *Salmonella* and the increased expression of over 50 proteins (Foster, 1991). Amongst these proteins are those that also offer some cross-protection against acid shock, heat shock, osmolarity and H_2O_2.

In addition *Salmonella* during the gastro-intestinal phase will have to face competition from the local micro flora for a limited amount of nutrients, low oxygen stress, IgA antibodies, bile salts and pancreatic enzymes. The physical presence of mucins may also protect the epithelial cells by preventing attachment of bacteria or contact with their toxins.

One of the major host defence mechanisms that the bacterium can come into contact with is the complement system. Once through the mucosal membrane exposure to complement can lead to opsonization or the lysis of the cells. Some bacteria have a polysaccharide capsule (e.g. *Salmonella* Typhi – Vi capsule) which can prevent this while others including Typhimurium depend on the complete O antigen of the lipopolysaccharide on the outer surface to resist complement action (Joiner, 1988). Other mechanisms include TraT surface exclusion protein encoded by *S.* Typhimurium virulence plasmid (Rhen, 1988) and the Rck protein which confers high levels of complement resistance to *S.* Typhimurium by preventing the membrane attack complex (MAC) from inserting into the membrane (Heffernan, 1992).

1.3.4 Resistance to phagocytes

There are several mechanisms by which bacteria can avoid being compromised by phagocytes such as macrophages. Firstly they can avoid being engulfed and thus stay in the extracellular environment. Should they be internalised they could escape from the phagosome or the fused phagosome-lysosome and then replicate within the 'safe' environment of the host cell cytoplasm, or they could prevent the fusion of the phagosome and lysosome. Finally they could persist inside the phagolysosome. *S.* Typhimurium demonstrates resistance to the oxidative and non-oxidative factors present in the macrophage (Alpuche-Aranda et al., 1994; Schwan et al., 2000).

Mutation studies involving the use of Tn10 insertions and subsequent screening for survival in murine peritoneal macrophages have revealed mutations that affect auxotrophy, response to oxidative stress, LPS, motility and colony morphology. These mutations led to a reduced level of survival and avirulence (Fields, 1986).

Oxidative or respiratory burst is one of the primary antibacterial functions of macrophages. It releases reactive oxidative intermediates (ROI's) such as superoxide, peroxide, hydroxyl radicals and hypochlorous acid. *S.* Typhimurium produces enzymes capable of neutralising ROI's including two catalases for peroxide and superoxide dismutase (Farr, 1991). LPS also gives protection from macrophage produced defensins (murine macrophages do not produce these), lysozyme, cathepsin G, elastase, azurocidin and other proteases. Mutants with incomplete LPS have been shown to be highly susceptible to granular extracts of human PMN lysosomal fractions (Rest, 1977).

1.4 Molecular pathogenesis of *Salmonella*

1.4.1 Plasmids.

Virulence plasmids have been found in a few of the serovars of *S. enterica* subsp. *enterica*, including Typhimurium. These are between 50 and 90 kB in size and are important in multiplication in the reticulo-endothelial system of the host. The plasmid of *Salmonella* Typhimurium (identified from the strain LT2) contains a 7.8 kB region containing the *spv* (*Salmonella* plasmid virulence) region which is needed for virulence. It contains 5 genes in a regulon, *spvRABCD*. The plasmid also encodes a fimbriae gene, the *pef* locus. Plasmids have also been shown to carry antibiotic resistances, such as those found in US strains of *S.* Newport and in the plasmid HCM1 from *S.* Typhi (Doublet et al., 2004).

1.4.2 Pathogenicity Islands.

Salmonella can contain a series of pathogenicity islands that play roles in invasion, survival and proliferation during interactions with the host cell. The presence of *Salmonella* Pathogenicity Island 1 (SPI-1) provides the capacity to invade while SPI-2 gives the capacity for systemic spread. The distribution of these two pathogenicity islands within *Salmonella* has shown an evolutionary development, both *S. bongori* and *S. enterica* possess SPI-1, whilst only *S. enterica* possesses SPI-2 (Ochman and Groisman, 1996). Pathogenicity islands and islets are identified by the lower G+C content of their regions compared to the rest of the *Salmonella* genome. Currently 14 have been identified, although their contribution to virulence is still to be fully assessed by mutational inactivation and complementation studies (Morgan, 2007).

SPI-1 is required by *Salmonella* for the invasion of host epithelial cells. It comprises a 40kB region encoding a type three secretion system (TTSS), effector proteins and

transcriptional regulators (Mills, 1995). Studies have shown that the expression of this region is controlled by environmental cues in the intestinal tract and the control genes *hilA* and *invF* (Lucas, 2001; Mills, 1995; Schechter, 1999). The secreted effector proteins are encoded on the island by; *avrA* (antiflammatory inhibition of NF_kB), *sipA* (preventing actin rearrangements), *sipB* (activating capsase 1, leading to apoptosis) and *sopE* and *sopE2* (activating Rho GTPases and cytoskeletal rearrangements). In addition the TTSS encoded by SPI1 can be used by proteins encoded by other pathogenicity islands.

SPI-2 has been associated with intra-macrophage survival and systemic disease in the host (Groisman and Ochman, 1997; Hensel, 2000). It contains 44 open reading frames (ORFs) including *ssrA/ssrB* and a TTSS. *In vitro* Ca^{2+} and Mg^{2+} ions are lowered within macrophages by SPI-2. *SsrA/ssrB* is a phosphor relay type sensor kinase. SPI-2 genes are needed for survival as they interfere in NADPH oxidase phagocyte dependent oxidative killing. SpiC is secreted into the macrophage cytosol and then prevents its fusion with the *Salmonella* containing phagosome.

SPI-3 is a 17 kb region, containing the *mgtCB* operon. This can confer growth in Mg^{2+} limited and intracellular environments. This is also involved in macrophage survival. Four other genes which encode cytoplasmic and putative membrane proteins are also present; *rmbA*, *misL*, *fidL* and *marT*. SPI-3 is regulated by a *phoP/phoQ* two component system. SPI-4 encodes a type 1 secretion system (T1SS) and the protein SiiE. SiiE is a giant non-fimbrial adhesin which is involved in the adhesion of *S.* Typhimurium to epithelial cells and whose synthesis and secretion co-incides with the activation of invasion genes (Gerlach, 2007). SPI-5 contains *pipABCD* and *orfX*. These are involved in host fluid and chloride secretion and a poor inflammatory response.

1.5 Virulence gene regulation

The two component PhoP/PhoQ system has been implicated in the regulation of aspects of *Salmonella* virulence. The system controls more than 40 genes. This system's functionality is required for virulence in mice and humans, survival within macrophages, growth on succinate as a sole carbon source and growth in the presence of limited magnesium (Groisman, 2001). The PhoQ is a sensor histidine kinase that phosphorylates the response regulator PhoP in response to environmental signals. PhoQ activity is repressed by cations of magnesium and calcium. Phosphoylated PhoP activates *pag* (PhoP-activated) genes such as non-specific acid phosphatase, calcium transporters and outer membrane proteins. PhoP-PO_4 represses genes from SPI-1 including *hilA* and genes encoding the type III secretion system. The two component system is highly induced within non-spacious phagosomes, where low pH and low concentrations of Mg^{2+} and Ca^{2+} trigger its activation. *PhoP* null mutants exhibit increased sensitivity to cationic antimicrobial peptides (CAMPs). Genetic analysis of *phoP* regulon resulted in identification of *pags* that mediate resistance to specific CAMPs by lipid A modifications (Groisman, 2001).

1.6 Drug resistance

Multidrug resistant (MDR) *S.* Typhimurium definitive type 104 (DT104) was first isolated and characterised in the UK in 1984 (Threlfall, 1994). These strains commonly had the penta resistance profile ACSSuT, which confers resistance to ampicillin, chloramphenicol, streptomycin, sulfonamides and tetracycline. This resistance has been shown to be chromosomally encoded (Sandvang, 1998) and therefore relatively stable. It is therefore likely that even reduction in the use of antibiotics and therefore their selective pressure would not necessarily lead to a loss of the resistance, as may be the case with plasmid encoded resistance. In 1990, 259 cases of MDR DT104 human infections were recorded in the UK. By

1996 this had risen to 4006 cases (Threlfall, 1997). This was a greater rate of increase then with *S.* Enteritidis in the previous decade (Bäumler, 2000). Internationally DT104 has also been on the increase, in 1992 8.7% of *S.* Typhimurium strains isolated were identified as DT104, by 2001 this proportion had increased to 33% (Helms, 2005).

In a study of 83 cases of MDR DT104 infection, 41% of patients (34) required hospitalisation and 10 of them subsequently died, although other illnesses may also have contributed (Wall, 1994). Of further concern has been the acquisition of additional resistance in DT104. From less than 2% in 1993, the proportion of isolates resistant to trimethoprim has risen to 24% whilst resistance to ciprofloxacin has increased from 1% to 14% in the same time period (Threlfall, 1997).

1.6.1 *Salmonella* Genomic Island 1

The resistance genes in MDR DT104 are contained in a 43 kb region known as the *Salmonella* Genomic Island 1 (SGI1). This region has a G+C content of 49.17% compared to between 51 and 53% with the rest of the Typhimurium genome (Boyd et al., 2001). The region lies between *thdF* and a novel retron sequence and consists of 44 open reading frames. The multidrug resistant region lies between ORF 28 and ORF 42. This region has been sequenced and shown to contain integrons conferring resistances to ampicllin and sulphonomides (bla_{PSE1} and *sul1*) and to streptomycin and spectinomycin (*aadA2*) with plasmid derived genes for chloramphenicol/florfenicol (*floR*) and tetracycline resistance (*tet(G)*) in between (Briggs, 1999; Lawson, 2004). Over time, different subclones of MDR DT104 with different resistance profiles have become more common (Threlfall, 2004).

The region outside of the multidrug resistance area has also been characterized. The genes from this region share similarity to plasmid genes involved in mating pair formation and DNA transfer. The putative products of ORF 5, ORF 11 and ORF 12 show similarity to the mating pair stabilization protein TrhN and pilus assembly proteins R0128 and TrhH from the IncH plasmid R27 found in *S*. Typhi (Boyd et al., 2001).

The region is not self-transmissible (Boyd et al., 2001). However it has been shown to be an integrative mobilizable element (Doublet et al., 2005a). The ORFS between DR-L and ORF 3 have been characterized as being a putative integrase and a putative excisionase.

1.7 Phage mediated evolution

The *Salmonella* genome shows the presence of many chromosomally bound bacteriophage like, or 'prophage', regions (Cooke et al., 2007). It has been postulated that bacteriophages evolve in a 'modular' manner, exchanging whole sequences of DNA between themselves, while maintaining the genes that facilitate their transfer. They can be incorporated in the bacterial chromosome via transduction or lysogenic conversion. The global rate of phage mediated genetic modification in bacteria has been estimated up to 20×10^{15} gene transfer events per second (Brüssow, 2004). Virulence factors known to be encoded on phages include; Shiga toxin, ADP-ribosyltransferase toxins, LPS-modifying enzymes and type III effector proteins.

Integrated bacteriophage regions have been identified in the *S*. Typhimurium genome (Cooke et al., 2007). They consist of 5 prophages; prophage 1(ST104), prophage 2(Gifsy-2), prophage 3, prophage 4 and prophage 5. Of these Gifsy-2 contains the genes; *sseI*, that codes for a type III effector protein, *sodC-1*, that codes for a superoxide dismutase and *gtgE*, which

has been hypothesised to have a role in virulence. The presence of phages within the bacterial genome is suggestive of a role they may have played in the development of the epidemic.

1.8 Comparative genetics

One of the drawbacks of traditional phenotypic characterisation of bacterial strains has been the difficulty in differentiating emerging infectious disease strains from others.. Grouping strains into phenotypically similar groups, i.e. through serotyping, does not necessarily reflect their genetic relatedness. Horizontal gene transfer of harmful genes, would generally be missed.

Comparative genomic hybridization (CGH), using microarrays, has become a powerful tool for the interrogation of differences between closely related bacterial species. The fine level of genetic detail provided by microarrays and the ability to compare entire genomes significantly increases the chances of discovering genetic variability (Joyce, 2002). *Salmonella* has been a target for CGH microarray (Chan, 2003; Porwollik, 2004) with comparisons made both between subspecies and within them. These studies have led to the identification of the core and variable components of *S. enterica* subspecies I (Anjum et al., 2005). The core invariant component of *Salmonella* genome shows an 83% overlap with the conserved homologues within *E. coli*, including genes involved with cell motility, metabolism, signal transduction, transcription and translation. This finding shows the common evolutionary pathways which connect *Salmonella* and *E. coli.*

The pathogenicity islands; SPI-1, SPI-2, SPI-3, SPI-4 and SPI-5, were shown to be conserved in all *S. enterica* subspecies I. Some variable regions identified are virulence determinants. Fimbriae, which are important in bacterial adherence to biotic and abiotic surfaces, are found throughout the strains. However only three fimbrial operons (*fim*, *bcf* and

stb) were conserved within the serovars examined. The CGH microarray method has therefore been demonstrated to be a valuable tool for considering the genetic make-up of *Salmonella* strains.

1.9 Conclusion and testable hypothesis

This brief introduction has given an overview of the biology and particularly the virulence factors of *Salmonella enterica*. Clearly alluded to, is the dynamic nature of the genome of *Salmonella enterica* with horizontal gene transfer by phage, plasmid and other mobile genetic elements playing a significant part in genome plasticity. Also cited above is the emergence of types that have come to the fore epidemiologically, which have a profound impact on human and animal health. It is interesting that types wax and wane to be replaced by the next emergent epidemic type. In simple terms, it may be envisioned that the emergence of a particular type may be due to a number of factors around the pathogen, the host and the environment. Given genome plasticity and our increasing knowledge of gene mobility, emergence may be through acquisition of new characteristics that provide some form of fitness advantage. There may be changes in host susceptibility such as through changing husbandry practices for animal management or immune naivety through subtle shifts in antigenicity of the infectious agent. In addition, there may be greater opportunities for infectivity through for example industrial food preparation and health care techniques, the global market and the rapid spread of pathogens across natural boundaries and climate change. In this study, I have attempted to find ways to explore possible reasons for the emergence of a particular type, namely *S.* Typhimurium DT104. Given the potential scope of such a study, I have focused down upon trying to ascertain what changes within DT104 may be associated with the epidemic. My testable hypothesis is that "DT104 emerged due to genetic changes that contributed to enhanced fitness". An interesting corollary to that is that

the decline of DT104 may also be due to loss or alteration to those genetic changes. Thus, my study was framed by the attempt to understand the genetics and phenotype of strains representing the epidemic cycle through investigating pre-, mid and post-epidemic strains of DT104.

2 Materials and methods

2.1 Bacterial strains

The bacterial strains used are shown in Table 2-1. These were all isolated from human and animal sources from the United Kingdom. These span the period from 1986 until the present. This encompasses the understood period of the DT104 epidemic, late 80's and throughout the 90's. They were isolated from individuals, both human and animal, suffering from salmonellosis. The animal strains came from species that come into contact with humans including bovine, porcine, avian and equine sources. Strains were sub-cultured from dorset egg slopes, provided by the Veterinary Laboratories Agency (VLA) and the Health Protection Agency (HPA), onto LB agar (Bertani, 1951) then further sub-cultured into LB containing 30% glycerol to be stored at -80^{o}C.

In addition seven sequenced control strains; *S. bongori* and the *Salmonella enterica* strains; *S.* Enteritidis phage type 4 (PT4), *S.* Typhimurium LT2, *S.* Typhimurium SL1344 (Wray, 1978) *S.* Gallinarum, *S.* Typhi CT18 and the sequenced *S.* Typhimurium DT104, are included.

2.2 Slide Agglutination

Strains were sub-cultured from glycerol stock into 3ml LB. These were then incubated at 37^{o}C overnight shaking at 150rpm. A 10µl loopful of phosphate buffered saline (PBS) was placed on a glass slide then a 10µl loop of the overnight culture to be tested was mixed in. Finally a 10µl loop of polyvalent O anti-sera was mixed in. Agglutination was identified by the formation of white 'grains' in the suspension. If a positive reaction to the polyvalent O anti-sera was seen then the procedure was repeated with the monovalent anti-sera that were

present in the polyvalent ant-sera that showed a positive response. By proceeding through a group of specific anti-sera the O antigen group can be identified. The process was repeated with a panel of H antigens (Jones, 2000).

2.3 Antibiotic Minimum Inhibitory Concentration

Following the validated British Society for Antimicrobial Chemotherapy (BSAC) method for minimum inhibitory concentration (Andrews, 2001) dilutions of 12 antibiotics (nalidixic acid, tetracycline, neomycin, ampicillin, ceftazidime, trimethoprim, chloramphenicol, gentamicin, streptomycin, cefotaxime, and ciprofloxacin) were prepared. Powdered antibiotic (Sigma) was weighed out and dissolved in the recommended solvent to a stock concentration of 256mg/L inside a safety cabinet. The solutions were then diluted in a doubling dilution, i.e. by taking 10ml of stock and adding to 10ml doubled distilled water a 128mg/L solution was made. This was continued until a range of dilutions from 128mg/L to 0.125mg/L was made. These were stored at 4^{o}C until used within the same 24 hour period.

10ml of each antibiotic preparation was then introduced into 10ml of molten IsoSensitest agar equivalent (ISA: Oxoid, Basingstoke, UK). These were then poured into 20ml plastic Petri dishes and left to set. The plates were then stored at 4^{o}C until needed.

Table 2-1 **Bacterial strains used for this study, including 32 *S*. Typhimurium DT104 test strains and the other *Salmonella* control strains included on the comparative genomic microarray**

ID	Reference	Serotype	Phage type	Isolation date	Resistance type *	Host
S01	P4722210	Typhimurium	DT104	1997	ACSSuSpT	Human
S02	P5416300	Typhimurium	DT104	2001	ACSSuSpT	Human
S03	P4082900	Typhimurium	DT104	1996	ACSSuSpT	Human
S04	P5289060	Typhimurium	DT104	2000	SSp	Human
S05	P5224010	Typhimurium	DT104	2000	ACSSuSpTTm	Human
S06	P5066840	Typhimurium	DT104	1999	ACSSuSpTTmNxCp	Human
S07	P5358880	Typhimurium	DT104	2001	ACSSuSpT	Human
S13	H0 4350 0377	Typhimurium	DT104	2004	-	Human
S15	H0 4208 0120	Typhimurium	DT104	2004	-	Human
S16	H0 4212 0222	Typhimurium	DT104	2005	-	Human
S19	H0 4286 0471	Typhimurium	DT104	2004	-	Human
S21	5252 0256	Typhimurium	DT104	2005	SSuSp	Human
S22	5302 0351	Typhimurium	DT104	2005	SSuSp	Human
S24	4322 0369	Typhimurium	DT104	2004	SSuSp	Human
S27	P097747/0	Typhimurium	DT104	1986	-	Human
S28	P097363/0	Typhimurium	DT104	1986	-	Human
S29	P097482/0	Typhimurium	DT104	1986	-	Human
S30	P097435/0	Typhimurium	DT104	1986	-	Human
A01	L01470-06	Typhimurium	DT104	2006	-	Environmental
A02	S06529-07	Typhimurium	DT104	2007	-	Equine
D01	S01760/03	Typhimurium	DT104	2003	ACSSuSpT	Porcine
D02	S05461/05	Typhimurium	DT104	2003	SSu	Porcine
D03	S02442/05	Typhimurium	DT104	2003	ACSSuSpT	Porcine
D04	B2436	Typhimurium	DT104	2005	ACSSuSpT	Poultry
D05	B2455	Typhimurium	DT104	2005	-	Turkey
D06	B2458	Typhimurium	DT104	2005	-	Bovine
D07	1341/96	Typhimurium	DT104	1996	ACSSuSpT	Poultry
D08	5085/98	Typhimurium	DT104	1998	ACSSuSpT	Poultry
D09	6703/00	Typhimurium	DT104	2000	ACSSuSpT	Bovine
D10	1041/96	Typhimurium	DT104	1996	ACSSuSpT	Porcine
D11	1424/98	Typhimurium	DT104	1998	ACSSuSpT	Equine
D12	3320/98	Typhimurium	DT104	1998	ACSSuSpT	Poultry
*S. b*ongori	ATCC 43975	*S. b*ongori	-	-	-	-
LT2	ATCC 700220	Typhimurium	-	-	-	-
SL1344	NCTC 13347	Typhimurium	-	-	-	-
DT104	NCTC 13348	Typhimurium	DT104	-	ACSSuSpT	-
PT4	NCTC 13349	Enteritidis	PT4	-	-	-
S. Gallinarum	NCTC 13346	Typhimurium	DT104	-	-	-
CT18	-	Typhi	-	-	ACSuTmT	-

* Ampicillin (A), Cholamphenicol/Floramphenicol (C), Ciprofloxacin (Cp), Nalidic Acid (Nx), Spectinomycin (Sp), Steptomycin (S), Sulphonamides (Su), Tetracycline (T), Trimethoprin (Tm)

Strains were sub-cultured from glycerol stock into 3ml LB. These were then incubated at 37^{o}C overnight and at 150rpm. The strains to be tested were diluted to 10^4 CFU/ml. The plates were dried for 30mins at 37^{o}C. Corrected culture were stamped with a multipoint inoculator onto the dried plates, these were left for 48 hours in an incubator at 37^{o}C.

2.4 Multilocus Sequence Typing (MLST)

All the test strains plus the sequenced DT104 strain underwent MLST screening. The strains were sub-cultured from glycerol stocks onto LB agar and incubated at 37°C overnight. DNA was isolated by harvesting the overnight cultures into 900µl of 0.1 M phosphate buffered saline (PBS). These were boiled at 100°C for 10mins before being centrifuged at 13,000 RPM (in a Microfuge) for 10mins. The supernatant was removed into sterile eppendorfs.

Primers for the 7 housekeeping genes from the *Salmonella* MLST database (http://web.mpiib-berlin.mpg.de/mlst) were acquired from Eurofin MWG Operon: *thrA* (aspartokinase, homoserine dehydrogenase), *purE* (phosphoribosylaminoimidazole carboxylase), *sucA* (α-ketoglutarate dehydrogenase), *hisD* (histidinol dehydrogenase), *aroC* (chorismate synthase), *hemD* (uroporphyrinogen III cosynthase) and *dnaN* (DNA polymerase III β subunit).

Polymerase chain reactions (PCR) were carried out for each sample for the 7 genes. 1µl of DNA lysate was used with 2µl of each respective primer and 45µl of Abgene mastermix containing ThermoPrime *Taq* DNA polymerase. The PCR cycling conditions are listed in Table 2-2 and the PCR primers are listed in Table 2-4. Reactions were carried out over 35 cycles on a GeneAmp® PCR System 9700, (Applied Biosystems).

Table 2-2 PCR cycle conditions for initial amplification of the 7 housekeeping genes

Temperature	Mins	No. Cycles
95°C	3.00	
94°C	0.20	
50°C	0.20	35
72°C	1.00	
72°C	5.00	
4°C	∞	

The initial products of the amplification were cleaned up using the Qiagen Mini Elute PCR Purification Kit (Qiagen, cat. no. 28006). Each product then had its DNA concentration estimated using a Labtech International Nano Drop ND-1000 spectrophotometer and run on an 1.5% agarose gel to check quality of the product. Those with clear bands of the expected size were then used for the sequencing reaction using BigDye Terminator Ready Reaction Mix v3.0 containing AmpliTaq® DNA Polymerase (Applied Biosystems), the conditions listed in Table 2-3 and the sequence primers listed in Table 2-5 were carried out over 30 cycles on a GeneAmp® PCR System 9700, (Applied Biosystems).

Table 2-3 Sequencing reaction cycle conditions for the cleaned up products of the initial PCR

Temperature	Mins	No. Cycles
96°C	3.00	
96°C	0.10	
		X
50°C	0.05	30
60°C	2.00	
4°C	∞	

Once the sequencing reaction was finished the reactions were stored at 4°C for a maximum of 24 hours. The reactions were cleaned up as follows. Each was transferred to a 1ml eppendorf. Independently 280µl of 3M sodium acetate (pH 4.6) was added to 7ml of 100% ethanol at room temperature. This was vortexed for 30 secs. 52µl of this mixture was added to each sample and vortexed. Samples were incubated in the dark at room temperature for 45mins. They were then centrifuged at 15,000 RPM at 4°C, the supernatant was discarded

and the pellets washed with 150µl of ice cold 70% ethanol. A further spin at 15,000 was carried out and the supernatant removed as before. Pellets were dried on the bench top at room temperature until all ethanol had evaporated, they were then sealed with nescofilm, packaged and sent for sequencing to the Sequencing Facility Department, University of Oxford.

Completed sequences were processed with the Lasergene SeqMan program from Dnastar, Inc. Sequences were proof read and trimmed to the correct fragment size. The completed sequences were then saved in a fasta format and submitted to the MLST database at the MPI für Infektionsbiologie website (http://web.mpiib-berlin.mpg.de/mlst/dbs/Senterica) to acquire the sequence type.

Table 2-4 Primer pairs used for the initial PCR amplification of internal fragments of the 7 houskeeping genes

Gene	Sequence	Product length
thrA: F	5'-GTCACGGTGATCGATCCGGT-3'	852 bp
thrA: R	5'-CACGATATTGATATTAGCCCG-3'	
purE: F	5'-ATGTCTTCCCGCAATAATCC-3'	510 bp
purE: R	5'-TCATAGCGTCCCCCGCGGATC-3'	
sucA: F	5'-AGCACCGAAGAGAAACGCTG-3'	643 bp
sucA: R	5'-GGTTGTTGATAACGATACGTAC-3'	
hisD: F	5'-GAAACGTTCCATTCCGCGCAGAC-3'	894 bp
hisD: R	5'-CTGAACGGTCATCCGTTTCTG-3'	
aroC: F	5'-CCTGGCACCTCGCGCTATAC-3'	826 bp
aroC: R	5'-CCACACACGGATCGTGGCG-3'	
hemD: F	5'-ATGAGTATTCTGATCACCCG-3'	666 bp
hemD: R	5'-ATCAGCGACCTTAATATCTTGCCA-3'	
dnaN: F	5'-ATGAAATTTACCGTTGAACGTGA-3'	833 bp
dnaN: R	5'-AATTTCTCATTCGAGAGGATTGC-3'	

Table 2-5 Primer pairs for sequencing reactions carried out on the products of the initial PCR

Gene	Sequence
thrA: sF	5'-ATCCCGGCCGATCACATGAT-3'
thrA: sR	5'-CTCCAGCAGCCCCTCTTTCAG-3'
purE: sF	5'-CGCATTATTCCGGCGCGTGT-3'
purE: sF1	5'-CGCAATAATCCGGCGCGTGT-3'
purE: sR	5'-CGCGGATCGGGATTTTCCAG-3'
purE: sR1	5'-GAACGCAAACTTGCTTCAT-3'
sucA: sF	5'-AGCACCGAAGAGAAACGCTG-3'
sucA: sR	5'-GGTTGTTGATAACGATACGTAC-3'
hisD: sF	5'-GTCGGTCTGTATATTCCCGG-3'
hisD: sR	5'-GGTAATCGCATCCACCAAATC-3'
aroC: sF	5'-GGCACCAGTATTGGCCTGCT-3'
aroC: sR	5'-CATATGCGCCACAATGTGTTG-3'
hemD: sF	5'-GTGGCCTGGAGTTTTCCACT-3'
hemD: sF1	5'-ATTCTGATCACCCGCCCCTC-3'
hemD: sR	5'-GACCAATAGCCGACAGCGTAG-3'
dnaN: sF	5'-CCGATTCTCGGTAACCTGCT-3'
dnaN: sR	5'-CCATCCACCAGCTTCGAGGT-3'

2.5 Comparative Genomic Hybridisation (CGH)

The previously indicated strains (Table 2-1) were compared using a whole genome array designed by the Sanger Institute, Cambridge (Thomson et al., 2004). It consists of specific unique PCR products, ranging in size from 200 – 500 bases from the *S.* Typhi CT18 genome, with the addition of specific PCR products from a further 6 strains representing different serovars. The control strains contribution to the array is outlined in Table 2-6. Positive and negative controls such as cy3 and genes fragments from the *Arabidopsis* genome were included on the slides. Each product was replicated at least 4 times on each individual slide.

Table 2-6 Source and of genes present of microarray slide.

Serovar / strain	Number of genes
S. Typhi CT18	4097
S. Gallinarum	83*
S. Typhimurium, LT2	397*
S. Typhimurium, SL1344	136*
S. Typhimurium, DT104	126*
S. Enteritidis, PT4	147*
S. bongori	539*

* Numbers indicate the unique genes from the six representative serotypes

The bacteria were cultured in LB broth overnight at 37°C and at 150 rpm in a shaking incubator. DNA was extracted using the Qiagen DNeasy® Blood & Tissue Kit (Qiagen cat no. 60504). 2µg of control DNA mix consisting of the 7 strains included on the array, and 2µg test DNA were labelled with fluorescent cyanine dyes Cy5 and Cy3 respectively, using the Invitrogen BioPrime DNA labelling System Kit (Invitrogen cat. no.18094-011). Labelled DNA was purified using the Qiagen MiniElute PCR Purification Kit (Qiagen cat. no. 28004). Dye incorporation was measured using a Labtech International Nano Drop ND-1000 spectrophotometer. A dye incorporation value of 80-100pmol/µl was required to continue with the assay.

The purified DNA was precipitated using 1/10th the volume of 3M NaAc pH 5.2 (SIGMA, cat. no. S7899) and 3 times the volume of room temperature 100% ethanol. The mixture was incubated at -70°C for 20mins to precipitate then centrifuged at room temperature for 10min at 13,000 rpm. The supernatant was discarded and 100µl of 70% ethanol was added. After a further centrifuge step of 5mins, the supernatant was removed by aspiration and air dried for 5mins. The DNA was resuspended in 30µl of hybridization buffer (5x sodium chloride sodium citrate (SSC), 6x Denhardt's reagent, 60mM TrisHCL pH 7.6, 0.12% sarkosyl; SIGMA; filter sterilized) and 6µl of tRNA (from yeast, SIGMA cat. no. R8508).

The Hybridization / DNA solution was heated at 95 °C for 2mins and cooled for approximately 10mins at room temperature. The hybridization buffer was then added to the middle of an inverted clean 60x25mm cover slip (VWR International cat. no. 631-0167). The microarray slide was gently lowered onto the cover slip to prevent bubbles and misplacement. The microarray slides were incubated in a hybridization chamber at 65 °C for 16-20 hours. The cover slip was removed by submerging in 65 °C washing buffer A (20xSSC, 20% sodium dodecyl sulphate (SDS); Sigma). The slide was then washed in washing buffer A for 2mins and transferred to washing buffer B (20xSSC; Sigma) for another 2mins, and washed twice. The slide was placed in 50 ml centrifuge tube and centrifuged at 1,500 rpm for 10mins.

Cy3/Cy5 fluorescent images of microarray slides were scanned by using Genepix 4000B scanner (Axon Instruments, Inc.). The obtained images were quantified with Genepix Pro software (Axon Instruments, Inc.) Signal intensities were corrected by subtracting the local background. Where only the control DNA has hybridized the Cy5 shows red, where both test and control DNA has hybridized a yellow spot appears.

Spots with a value for median signal intensity minus the background below 50 (in both the control and test channel) were discarded due to low signal intensity that could lead to ambiguous results. Per spot per chip intensity-dependent Lowess normalization was performed to compensate for unequal dye incorporation. The ln(Cy3/Cy5) ratio intensity of all spots from the hybridization was calculated and the mean ln(Cy3/Cy5) from up to eight data points per gene (i.e. two slides with four repeats per slide), was used in all subsequent data analyses. The relationship between strains was calculated by average-linkage hierarchical clustering using the Pearson Correlation as part of the GeneSpring microarray analysis software V5.0 (Silicon Genetics). Exact cut-offs for present/absent genes are often inaccurate for microarray data, therefore the presence/conserved or absence/divergent nature of each gene was calculated by the final calculation ln(Cy5/Cy3) ratio intensity of the normalized data. The present/absent genes were displayed in GeneSpring by giving a standard GeneSpring coloured representation. Those genes that can be considered to be absent/divergent in this experiment with the greatest degree of certainty lie at the blue end of the scale, around an ln(Cy5/Cy3) ratio of 0. Those genes that are present/conserved with the greatest degree of certainty lie in the yellow region and have a ln(Cy5/Cy3) ratio of around 1. The red end of the scale indicates an increased hybridization with the test strain and may be to due to presence of genes with multiple copies.

2.6 PCR for artA/B presence in strains

Using the identified *artA/B* toxin region in DT104 (Saitoh, 2005) and the sequence of DT104 available (Dougan) PCR primer pairs were designed for *S.* Typhimurium DT104 genes *artA*, *artB*, and a bridging region covering *artA* and *artB*. The PCR cycling conditions used are listed in Table 2-7. PCR primer pairs are listed in Table 2-8. DNA was extracted using the Qiagen DNeasy® Blood & Tissue Kit (Qiagen cat no. 60504).

Table 2-7 PCR cycle conditions for amplification of the ArtA/B region

Temperature	Mins	No. Cycles
95°C	3.00	
94°C	0.20	
50°C	0.20	35
72°C	1.00	
72°C	5.00	
4°C	∞	

2.7 Sequencing

To investigate possible variations with the *artA/B* region a sequencing study was carried out. Using the identified *artA/B* toxin region in DT104 (Saitoh, 2005) and the sequence of DT104 available (Dougan) primers were designed to encompass the entirety of the *artA* and *artB* genes, giving a predicted product size of 1380bp (Table 2-8). DNA was extracted using the Qiagen DNeasy® Blood & Tissue Kit (Qiagen cat no. 60504). The PCR cycling conditions were used as in Table 2-7. Sequencing primers are listed in Table 2-8. PCR products were sent to the Veterinary Laboratories Agency sequencing department for processing. Returned sequences were processed with SeqMan and analysed with MegAlign, both programs from the DNASTAR Lasergene suite of applications.

Table 2-8 PCR primer pairs for amplification of areas within the pertussis toxin region

Gene	Sequence	Product length
artA/B bridge F	5'-GTAAGTACCCGGGAAAGGATT -3'	417 bp
artA/B bridge R	5'- ATACACGCCATAAGACAGATTG-3'	
artA F	5'-GAGACCTCCGGATGTGATTTTC -3'	465 bp
artA R	5'-ACCCCAGGATTTGATTGTGTA -3'	
artB F	5'-GCTCTTACTCTTGCGTCGTT -3'	388 bp
artB R	5'-TAGGTCCCATACAATAATCT -3'	
artA/B seq F	5'-GGTGGCTTGCCTGATAGTT-3'	1380 bp
artA/B seq R	5'-ACATAATCGTCTCACCAACAAG-3'	

2.8 Phenotypic MicroArray™.

The growth phenotypes of the DT104 test strains were assessed using the Phenotypic MicroArray™ system from Biolog, Inc., utilising plates PM1, PM2a, PM3b and PM4a. These are 96 well microtitre plates each containing a defined medium with the addition of sole carbon sources (PM1 and PM2a), sole nitrogen sources (PM3b) or sole phosphorus and sulphur sources (PM4a). In addition, to each well was added tetrazolium violet which is an indicator of respiration when it is reduced. PM1 and PM2a contain 190 carbon sources (with ammonium as a nitrogen source), PM3b contains 95 individual nitrogen sources and PM4a contains 59 phosphorus and 35 sulphur sources. To these last two plates succinate was added as a carbon source, with the inoculating fluid (see below).

Firstly a carbon solution stock of 2M sodium succinate and 200 μM ferric citrate (pH 7) was prepared (Table 2-9). The solution was sterilized by filtering through a 0.8 μm filter and stored at in 5 ml aliquots at 4^{o}C. 10 ml of IF-0 (inoculating fluid, no carbon, Biolog, Inc.) was aliquoted into a 150 mm sterile capped test tube for each strain being prepared. The strains were grown on a LB agar plate by streaking the bacterium from isolated colonies and allowing it to grow overnight at 37°C. If the cells were streaked from a frozen glycerol stock, they were sub-cultured a second time. Cells were taken from the LB agar plate using a sterile swab and transferred into the sterile capped tube containing 10ml of IF-0. The cell suspension was stirred with the swab and then inverted to obtain a uniform suspension. The turbidity of the suspension was checked with the Biolog, Inc Turbidimeter and cells or IF-0 added to achieve 85% T (transmittance).

Table 2-9 Preparation of carbon source for Phenotypic MicroArray, for plates PM3b and PM4a.

Carbon Source	Formula Weight	Per 100 ml
2 M sodium succinate	270.1	54.02 g
200 µM ferric citrate	244.9	4.9 mg

Independently the inoculating fluids used for inoculating phenotype array panels were prepared as follows (Table 2-10). 100µl of the relevant inoculating fluid was added to each well of the plates. The inoculated plates were then placed inside the OmniLog™ incubator and incubated for 30 hours at 37°C. The OmniLog™ software recorded the level of colour change every 15mins. Kinetic data was analyzed with the OmniLog™ – Pm software version OL_PM_FM/Kin 1.30, and imported into Microsoft Excel.

Table 2-10 Preparation of inoculating fluid for Biolog plates.

Ingredient	PM 1 – 2a	PM 3b – 4a
IF-Base	10 ml **(IF0A)**	10 ml **(IF0A)**
Dye Mix A 100X	0.12 ml	0.12 ml
Carbon Source (200mM)	-	0.12 ml
cell density 85%T	0.12 ml	0.12 ml
Thiamine*	0.003 ml	0.003 ml
Sterile water	1.76 ml	1.64 ml
Final Volume	12 ml	12 ml

2.9 Growth rate assay

The growth of the strains in minimal medium was measured by reading the OD_{600} every 30mins over 240 cycles, for a total of 120 hours, using the FLUOstar Optima automated microplate based multi-detection reader (BMG Labtechnologies, Offenburg, Germany). *S.*Typhimurium LT2 and *S.* Gallinarum were included as control strains.

Inoculating cultures were grown from glycerol stocks into 3ml LB broth, at 37^{o}C shaking at 150 rpm overnight. MOPS minimal medium (Neidhardt, 1974) was used in the preparation of the test media (Table 2-11). The standard assay uses 0.4% glucose (w/v) final concentration and this was replaced by compounds highlighted by the Phenotypic MicroArray study, either 20% glyoxylate or tartaric acid; in addition LT2 grown in glucose was used as a reference strain.

Table 2-11 Inoculating fluid preparation for FLUOstar assay.

Medium A		**Medium B**	
Reagent	**Volume / µl**	**Reagent**	**Volume / µl**
MOPS	1000	MOPS	1000
K_2HPO_4	100	K_2HPO_4	100
H2O	8700	H2O	8700
20% Glucose	200	20% Glyoxylate or Tartaric acid	200
Total	10000	Total	10000

1498.5µl of chosen medium for each strain to be tested was aliquoted into labelled eppendorf tubes. To this 1.5µl of overnight culture was added and was mixed by pipetting to give ≈ 1x10^5 cells. 200µl of the medium/strain mix was added to at least 3 wells of an 96 well plate from IWAKI®. This was then placed in the FLUOStar Optima reader.

Growth of the strains was measured for glyoxylate concentrations of 0.2, 0.4, 0.8% (w/v) final concentration and with tartaric acid at 0.1, 0.2, 0.4, 0.8, 1.6, and 3.2% (w/v) final conc. These concentrations encompass those present on the Phenotypic MicroArray plates. The FLUOstar machine took OD readings every 900secs for 150–240 cycles, to cover the growth up to stationary phase. Data on growth is outputted into an Excel spreadsheet for processing.

2.10 Macrophage persistence study

A tissue culture assay was carried out to measure the persistence of some of the strains inside macrophages. Preliminary tests were carried out with two cell lines. The Murine RAW 264.7 and human U937 macrophages were tested. Initial optimisation showed more consistant results with the human cell line and therefore this line was chosen to move forward with. 72 hours prior to infection cells were seeded into 24 well plates at $2x10^5$ cells ml^{-1}. The strains chosen were sub-cultured into 20ml LB broth and left shaking at 150rpm at 37°C overnight. Overnight cultures were centrifuged at 13,000 rpm for 10mins, washed with 10ml 0.1M PBS and then re-spun at 13,000 rpm. Washed cultures were re-suspended and corrected to an OD_{600} of 1.2 with the addition of further PBS.

Dulbecco's Modified Eagle Medium (DMEM, Sigma) culture medium with 1% L-Glutamine (Sigma) and 1% essential amino acids (Sigma) was prepared. Bacterial suspension was added to prepared culture medium (500µl to 9.5ml of medium) to give a final concentration of bacteria of $\approx 2x10^7$ colony forming units (c.f.u.) ml^{-1}. 1ml of this suspension was added to the tissue culture wells in triplicate, giving a multiplicity of infection ratio of 1:100. The plates were then left at 37°C in a 5% CO_2 incubator for 2 hours.

After 2 hours one plate was washed 2x with Hanks Balanced Salt Solution (BSS, Sigma) then the cells were lysed with 0.1M PBS containing 1% Triton (Sigma). Supernatant from this plate then underwent a dilution series and was plated out for colony counting. The other plates were washed 3x with Hanks BSS. 1ml of fresh culture medium containing $100\mu g/ml^{-1}$ of gentamicin was then added to each well, and the plates left incubating. Plates were when removed at different time points (5 hours or 24 hours). They were then washed

and lysed with 0.1M PBS containing 1% Triton (Sigma). Supernatant from this plate then underwent a dilution series and was plated out for colony counting.

3 Characterisation of strains

The bacterial test and control strains chosen for this study are listed in the method chapter in Table 2-1. The test strains were isolated from humans suffering from salmonellosis (18 strains) and from animal hosts (14 strains) from the United Kingdom. They span the period from 1986 through to 2007. This time frame encompasses the understood period of the *S.* Typhimurium DT104 epidemic, from the late 1980's and through the 1990's. Therefore, this strain selection should permit the investigation of genetic changes that may have taken place during this period.

As well as time and host species, another factor taken into account for strain selection was the antibiotic resistance profile. Resistance in DT104 was first reported at the beginning of the epidemic period (Threlfall, 1994). Strains chosen from the pre-epidemic period were generally sensitive to antibiotics. Strains from the epidemic period have the 'typical' penta resistance profile ACSSuT, however, variant resistance profiles as well as fully sensitive profiles were also observed during the epidemic. A selection of strains reflecting this diversity was chosen. In the post-epidemic period, strains showed both sensitivity and resistance and so a further selection of strains was made from this period. Overall 20 of the strains showed differing degrees of antibiotic resistance while 12 were completely sensitive. The previous paragraph posed a testable hypothesis that the strains arising from the epidemic were likely to show genetic changes probably associated with the rise and fall of the epidemic. The evidence of varying resistances from fully sensitive, to resistant, to sensitive again was preliminary evidence to suggest the hypothesis was potentially sound.

The test strains were provided by the Health Protection Agency (HPA) and the Veterinary Laboratories Agency (VLA) as cultures on Dorset egg slopes. They were sub-

cultured onto the nutrient rich Luria-Bertani (LB) agar (Bertani, 1951) then further sub-cultured in triplicate into LB broth containing 30% glycerol to be stored frozen at -80^{o}C for long term storage. An issue that had to be considered was the effect that prolonged storage at low temperature and repeated thaw and re-freezing could have on the genetic make-up of the strains. Bacterial survival processes may lead to genetic changes that would mask or replace any differences that may exist in the initial isolates. Given this concern, it was important to demonstrate that strains selected for study still conformed to the data (serotype, phagetype and resistance profile) ascribed to each strain on initial storage.

Furthermore, additional genetic tests such as PFGE and MLST can be used to assess relatedness. Relatedness of course assumes that we are dealing largely with a single clonal type that has acquired resistances and possibly other genetic changes over the period of the epidemic.

In addition, seven control strains were provided by the VLA and the Wellcome Trust Sanger Institute, Welcome Trust Genome Campus, Hinxton, Cambridge. These were: *S. bongori* and the *Salmonella enterica* strains; *S.* Enteritidis phage type 4 (PT4), the *S.* Typhimurium strains LT2, SL1344 and DT104, *S.* Gallinarum and *S.* Typhi CT18. These were sub-cultured and stored in the same manner as the field isolates. The seven control strains have had their genomes sequenced, and this information and their annotations are available from GenBank and the Sanger Institute website. *S. bongori* ATCC43975 (Sanger Institute) (Dougan), *S.* Enteritidis P125109 (GenBank AM933172) (Thomson et al., 2008), *S.* Typhimurium DT104 NCTC13348 (Sanger Institute) (Dougan), *S.* Typhimurium LT2 (GenBank AE006468) (McClelland et al., 2001), *S.* Typhimurium SL1344 NCTC13347 (Sanger Institute) (Dougan) *S.* Gallinarum 297/91 (GenBank AM933173) (Thomson et al., 2008), *S.* Typhi CT18 (GenBank AL513382) (Parkhill et al., 2001).

3.1 Preliminary characterisation

3.1.1 Serotyping

To confirm the designation of the test strains as *Salmonella* Typhimurium DT104 slide agglutination and phage typing were carried out. All test strains were confirmed as being *Salmonella* Typhimurium via slide agglutination, with the Kauffmann-White serotype 1,4,[5],12;i;1,2. The O-antigens O4, O5 and O12, and phase 1 H-antigen i and phase two H-antigens 1 and 2 were detected. Phage typing was carried out in conjunction with the VLA *Salmonella* reference laboratory within the Food and Environmental Safety Department, and the strains were confirmed as being definitive type 104. Both these characterisations were repeated periodically throughout the study to confirm that no contamination had occurred.

3.1.2 Antibiotic resistance

The evolution of antibiotic resistance within Gram negative bacteria has become an issue of increasing importance, with implications for the treatment of disease in animals and humans. Numerous tests and methodologies have been developed to ascertain quickly and apply quantitative values to such resistances. Therefore the antibiotic profiles associated with the test strains were confirmed with an antibiotic MIC following the scheme developed by the British Society for Antimicrobial Chemotherapy (BSAC). The thresholds acquired from the test fell within the expected values provided by the 'BSAC Methods for Antimicrobial Susceptibility Testing' Version 5, February 2006' (Andrews, 2006). The results, displayed in Table 3-1, confirmed the antibiotic resistance profiles of the test strains that had been assigned to them before the study Table 2-1.

3.2 Further characterisation

3.2.1 Multilocus Sequencing Typing (MLST)

Multilocus sequence typing was used to discriminate between strains by looking at the changes in bases within stable chromosomal housekeeping genes, as it allows an insight into the evolutionary development of strains and their relationships to each other. The MLST scheme for *S.* Typhimurium takes advantage of 7 housekeeping genes (Anderson, 1977) (see methods) and has classified the strains as shown below in Table 3-2 and in Figure 3-1.

The DT104 strains were generally assigned the sequence type ST19. The exceptions to this being stains S15 and A02, which had the sequence type ST36 and strains A01 and D02 which had independent unknown sequence types.

The database of submitted strain information at the Max-Planck-Institut für Infektionsbiologie website (http://web.mpiib-berlin.mpg.de/mlst/dbs/Senterica) shows 230 Typhimurium strains with the associated ST19 type and only 3 with the ST36. Of the listed ST19 strains; 116 have no phage type assigned, 45 are DT104, 7 are DT193, 12 are DT204, 10 are DT99, 4 are DT49 and there are miscellaneous others.

Table 3-1 **Antibiotic minimum inhibitory concentrations for the 30 test strains for a panel of 12 antibiotics. The BSAC standards for each antibiotic is listed at the top of the table, and the level at which no further growth was observed for each of the test strains is listed below. If the value at which no growth was noted is greater that the BSAC standard then the strain was considered 'resistant' to that antibiotic**

	Antibiotic mg/L											
	Nalidixic Acid	Streptomycin	Tetracycline	Neomycin	Gentamicin	Sulfadiazine	Ampicillin	Ceftotaxime	Ceftazidime	Chlorophenicol	Trimethoprim	Ciprofloxacin
BSAC standard	*>128	8	1	-	4	32	16	2	8	8	0.12	0.015
Strain												
S01	128	>128	>128	4	1	>128	>128	0.5	8	>128	0.25	<0.125
S02	128	>128	>128	8	2	>128	>128	0.5	4	>128	0.25	<0.125
S03	128	128	>128	4	1	>128	>128	0.25	8	>128	0.25	<0.125
S04	128	128	4	4	1	2	2	0.25	4	16	0.125	<0.125
S05	128	128	>128	4	1	>128	>128	0.25	4	128	16	<0.125
S06	>128	>128	>128	4	1	>128	>128	0.5	4	>128	8	0.25
S07	128	128	>128	4	1	128	>128	0.25	4	>128	0.25	<0.125
S13	128	16	4	4	1	16	4	0.5	4	4	0.125	<0.125
S15	128	16	4	4	1	16	1	0.25	4	4	0.125	<0.125
S16	128	32	4	4	1	8	2	0.5	4	4	0.25	<0.125
S19	128	16	4	4	1	8	4	0.5	4	4	0.125	<0.125
S21	128	>128	4	4	1	>128	2	0.5	4	4	0.125	<0.125
S22	128	>128	4	8	1	>128	2	0.5	4	8	0.125	<0.125
S24	128	>128	4	4	1	>128	4	0.25	4	4	0.25	<0.125
S27	128	32	4	8	2	16	4	0.5	4	4	0.125	<0.125
S28	128	32	4	4	1	16	2	0.25	4	4	0.25	<0.125
S29	128	16	4	4	1	8	4	0.5	4	4	0.25	<0.125
S30	128	32	4	4	4	16	2	0.5	4	8	0.125	<0.125
A01	128	2	4	4	1	4	2	0.25	4	4	0.125	<0.125
A02	128	2	4	4	1	4	2	0.25	4	4	0.125	<0.125
D01	128	>128	>128	4	1	>128	128	0.25	4	>128	0.125	<0.125
D02	128	>128	4	4	1	>128	4	0.25	4	4	0.125	<0.125
D03	128	>128	>128	4	1	128	>128	0.5	4	>128	0.125	<0.125
D04	128	>128	>128	4	1	128	>128	0.25	4	>128	0.125	<0.125
D05	128	2	2	4	1	4	2	0.25	4	4	0.125	<0.125
D06	128	1	4	4	1	8	2	0.25	4	4	0.125	<0.125
D07	128	>128	>128	8	1	>128	>128	0.5	4	>128	0.25	<0.125
D08	128	>128	>128	4	1	>128	>128	0.25	4	>128	0.25	<0.125
D09	128	>128	>128	4	1	>128	>128	0.25	4	>128	0.125	<0.125
D10	128	>128	>128	4	1	128	>128	0.25	4	>128	0.125	<0.125
D11	128	>128	>128	4	1	>128	>128	0.25	4	>128	0.125	<0.125
D12	128	>128	>128	4	1	>128	>128	0.5	4	>128	0.125	<0.125

The comparative rareness of strains with the ST36 type may be an indication of either their infrequent isolation, perhaps due to a different host reservoir or a lower incidence of clinical symptoms, or may be indicative of their more recent emergence. The same rationalization can be used for unknown ST types of strains A01 and D02.

An interesting observation was the two strains A01 and D02 that had an unknown MLST ST type were similar to ST19 with only a single allele, *purE*, harbouring one single-nucleotide polymorphism (SNP) difference. It would be unwise to draw any conclusions about recent divergence as studies from this laboratory (Pan et al., 2009) showed that a single *purE* SNP was associated with major genomic changes (loss of an ABC transporter system) associated with PT11 of *S.* Enteritidis. This was a significant differential from *S.* Enteritidis PT4 and the other 'virulent' types. It is, however, tempting to suggest that A01 and D01 are likely to be more closely related to ST19 than the ST36 type that has 4 allele differences.

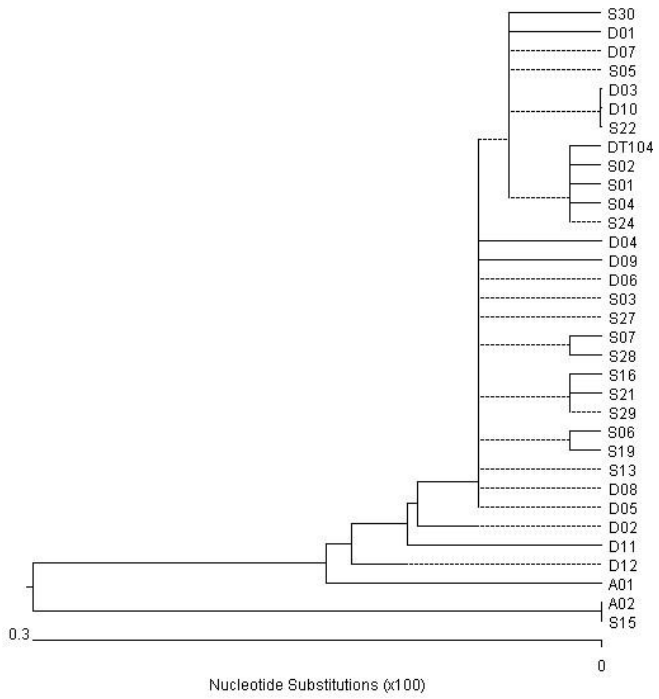
Figure 3-1 shows a phylogenetic tree generated by the ClustalW program from the DT104 MLST sequences. Generally the tree shows that strains A02 and S15 group furthest away, the next group contains animal strains A01, D02, D11 and D12. The rest of the strains show very little difference, although some strains show a higher degree of similarity with S22, D03 and D10 being very alike and S01, S02, S04 and S24 grouping with the sequenced DT104 strain. Overall, excluding S15 and A02, the DT104 strains seem to be very closely related when considering these 7 genes.

Table 3-2 **The sequence types obtained by the MLST for the 32 test strains. 7 housekeeping genes were the targets of the MLST primers and their obtained sequences were submitted to the MPI MLST database for identification**

Strain	MLST primer gene target							ST
	aroC	*dnaN*	*hemD*	*hisD*	*purE*	*sucA*	*thrA*	
DT104	10	7	12	9	5	9	2	ST19
S01	10	7	12	9	5	9	2	ST19
S02	10	7	12	9	5	9	2	ST19
S03	10	7	12	9	5	9	2	ST19
S04	10	7	12	9	5	9	2	ST19
S05	10	7	12	9	5	9	2	ST19
S06	10	7	12	9	5	9	2	ST19
S07	10	7	12	9	5	9	2	ST19
S13	10	7	12	9	5	9	2	ST19
S15	18	14	12	9	5	18	21	**ST36**
S16	10	7	12	9	5	9	2	ST19
S19	10	7	12	9	5	9	2	ST19
S21	10	7	12	9	5	9	2	ST19
S22	10	7	12	9	5	9	2	ST19
S24	10	7	12	9	5	9	2	ST19
S27	10	7	12	9	5	9	2	ST19
S28	10	7	12	9	5	9	2	ST19
S29	10	7	12	9	5	9	2	ST19
S30	10	7	12	9	5	9	2	ST19
A01	10	7	12	9	71	9	2	**UNKNOWN**
A02	18	14	12	9	5	18	21	**ST36**
D01	10	7	12	9	5	9	2	ST19
D02	10	7	12	9	71	9	2	**UNKNOWN**
D03	10	7	12	9	5	9	2	ST19
D04	10	7	12	9	5	9	2	ST19
D05	10	7	12	9	5	9	2	ST19
D06	10	7	12	9	5	9	2	ST19
D07	10	7	12	9	5	9	2	ST19
D08	10	7	12	9	5	9	2	ST19
D09	10	7	12	9	5	9	2	ST19
D10	10	7	12	9	5	9	2	ST19
D11	10	7	12	9	5	9	2	ST19
D12	10	7	12	9	5	9	2	ST19

All strains, except S15 shown to have sequence type ST19. Strains S15 and A02 shown to have sequence type ST36, strains A01 and D02 shown to have an unknown sequence type. Sequence types retrieved from http://web.mpiib-berlin.mpg.de/mlst/dbs/Senterica database

Figure 3-1 **Phylogenetic tree showing the relatedness of the 32 test strains and the control DT104. Created by comparing the sequences of the 7 housekeeping genes of the MLST using the ClustalW program**



For bacteria, a mutation frequency (the frequency in which an error in DNA replication results in change in an individual base pair) can lie between 10^{-7} and 10^{-11} per base pair per single round of replication (Madigan et al., 2006). Thus for an average gene of around 1000 base pairs, mutations can occur in the range of 10^{-4} to 10^{-8} per generation. The low values from the MLST data, with A02 and S15 showing 0.3 substitutions per 100 bases, shows that the variation within these metabolic genes is very low. This suggests that those strains which diverge from the main group are still evolutionarily closely related.

3.2.2 Pulsed field gel electrophoresis (PFGE)

Claudia Lucarelli from the Instituto Superiore di Sanita (ISS) in Italy carried out a PFGE assay, using the Xba 1 restriction endonuclease, on the 18 strains from human hosts, and has kindly allowed the results to be used. The results are shown in Figure 3-2. They show the clonal nature of these strains. Each strain had the same PFGE profile, XB0061, except for strain S15 which had the profile XB0088.

Comparing to the SalmGene PFGE database (HPA, http://www.hpa-bionum.org.uk/bionumerics/salm_gene/) XB0061 shows a common profile with other DT104's, only two other strains show the XB0088 profile and these both are phage type U277. This phage type has been isolated in Norway from passerine birds and has been shown to fall into the same PFGE group as some DT99 strains (Refsum, 2002).

Discussion

The classification results of the test strain have confirmed that all are *S.* Typhimurium and have the phage type DT104. In addition they have established that the resistance profiles associated with the strains before the study are correct. It was therefore confidently assumed that the storage of the strains at -80^{o}C has not significantly affected them and that they will represent the genetic content they had when isolated and first stored

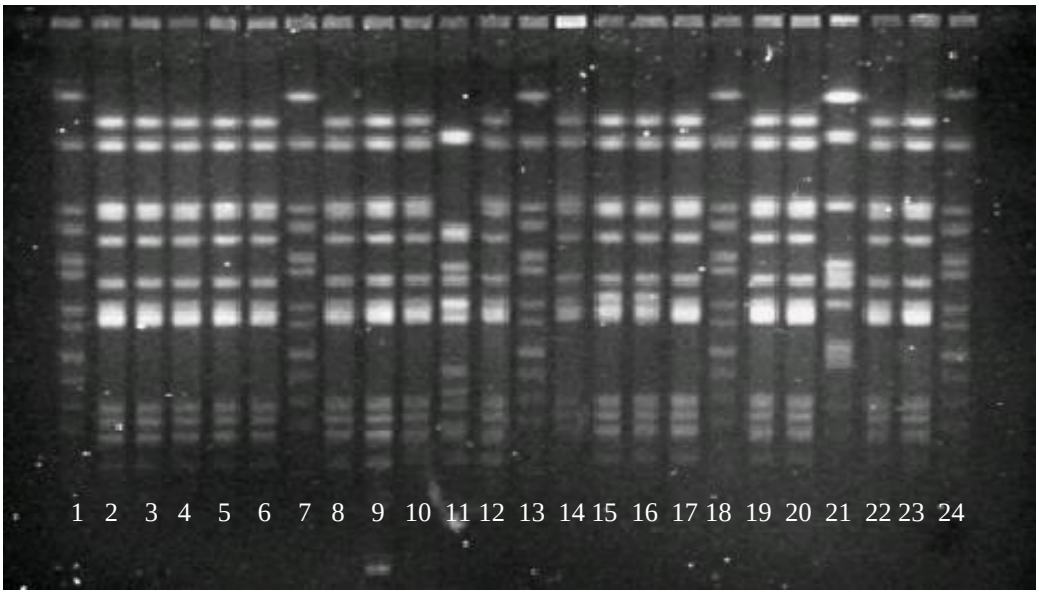
Further classification by MLST and PFGE was carried out to differentiate between strains. The use of pulsed-field gel electrophoresis to differentiate between bacterial strains has become the international 'gold standard' for epidemiology. It allows the direct comparison of strains by their banding patterns. The PFGE study

carried out on the human DT104 strains showed that apart from one of the strains, S15, all showed the same profile.

The MLST data showed that generally the DT104 strains were clonally grouped into ST19, but also revealed two strains with ST36 and two with an unknown ST type. The fact that the ST36 group differs by more SNP's then the unknown group may suggest that the later was closer related to the ST19 'core' group. However even small changes can have a profound effect on the bacteria as has been shown in *S.* Enteritidis (Pan et al., 2009).

The use of seven housekeeping genes as the target of the MLST does have limitations. Narrowing just to those genes precludes the detection of variation in genes that may have more significance to bacterial epidemicity, such as virulence genes. The stability of central metabolic genes also reduces the discriminatory power of the test when considering very closely related strains. The general clonal nature of the strains has been demonstrated by the MLST and PFGE studies. This is suggestive that epidemicity with DT104 is due to more discreet changes rather than large scale evolutionary development.

Figure 3-2 **Electrophoresis gel showing PFGE results for human strains S01 – S07, S13, S15, S16 and S19, S21, S22, S24 and S27 – S30. All samples show the typical DT104 pattern XB0061 profile except S15 (lane 11) which shows the XB0088 profile.**



Lane	Sample	Profile
1	Ladder	N/A
2	S01	XB0061
3	S02	XB0061
4	S03	XB0061
5	S04	XB0061
6	S05	XB0061
7	Ladder	N/A
8	S06	XB0061
9	S07	XB0061
10	S13	XB0061
11	S15	XB0088
12	S16	XB0061
13	Ladder	N/A
14	S19	XB0061
15	S21	XB0061
16	S22	XB0061
17	S24	XB0061
18	Ladder	N/A
19	S27	XB0061
20	S28	XB0061
21	Ladder	N/A
22	S29	XB0061
23	S30	XB0061
24	Ladder	N/A

Results courtesy of C. Lucarelli (ISS).

4 Comparative genetic analysis

Introduction

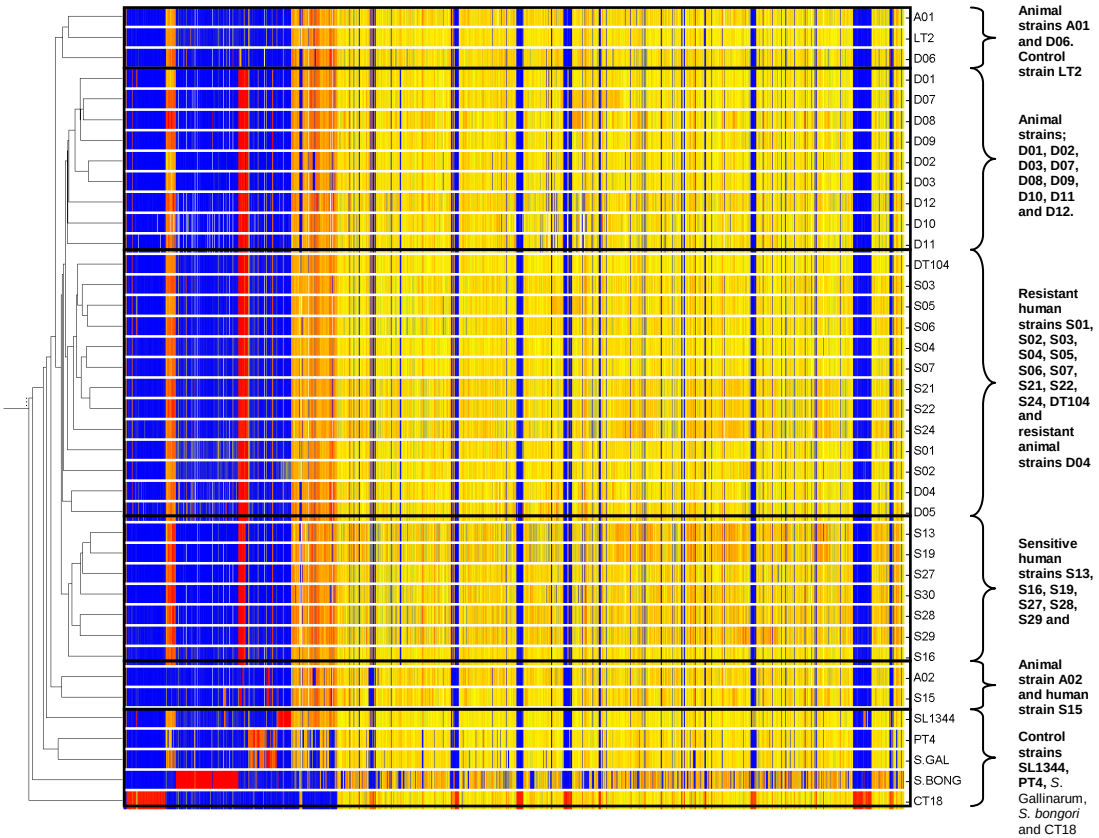
The previous section, described the criteria for strain selection and the initial studies that aimed at verifying two key points. First, that the strains used conformed to their initial description upon storage and, second, that there was little or no evidence to suggest that storage was associated with major genetic changes *per se*. The data generated gave confidence that the strains were as near to their original state on storage as could be anticipated: there may of course be some changes such as point mutations and possibly rearrangements and deletions.

To gain an understanding of the possible changes in the panel of strains at the whole genome level, it was decided to undertake comparative genomic analysis following the methods first described by Anjum et al., (2005). This approach was selected as a relatively rapid way of determining the presence or absence of genes against a pan-*Salmonella* array. The outcome of such an approach would be to identify any deletions associated with prolonged storage which one could anticipate as being possibly random events present in individual strains. Conversely, one might anticipate any loss or acquisition associated with the emergence of the epidemic strains may be seen as consistent changes in phenotypically related strains. This section describes comparative genomic analyses.

4.1 CGH Microarray results.

The comparative genomic microarray was carried out as outlined in the methods. The data tables are included with the electronic submission. The resultant data was analysed with the Genespring software and the results of a Pearson correlation clustering is shown in Figure 4-1. This groups the control strains as most divergent, separates the antibiotic resistant strains from the sensitive strains and further divides the animal resistant strains from the human resistant strains and the sensitive strains into three distinct groups.

Figure 4-1 **Grouping of all test and control strains for all genes on the CGH array using the GeneSpring Software. Clustered using Pearson correlation. Presence of genes shown by either yellow or red colour, blue indicates absence of the gene in that strain**



4.2 General comparison to sequenced DT104 strain

Comparing the samples to the genes present in the sequenced epidemic DT104 strain allows a more detailed study of the genetic differences between strains. This comparison suggests splitting the strains into 5 groups; human resistant strains, animal resistant strains (1), animal resistant strains (2), sensitive strains (1) and sensitive strains (2). See Table 4-1.

The first identified group contains the human antibiotic resistant strains (S01-S07, S21, S22 and S24) which have a high similarity (between 99.29% and 99.57%) to DT104. Some genes missing from multiple strains, when compared to the sequenced DT104, include; phage related genes (SBG0906, SEN1384 and SYY4610), SEN4371 - *pefA* – fimbrial polypeptide (in strains S05, S06, S07, S21, S22 and S24), *rsdB* – resolvase (in strains S05 and S06), STY0220 – *fhuD,* ferrichrome-binding periplasmic protein precursor (in strains S21 and S22) and STY3961 - *ccmC2* cytochrome related protein (in strains S06 and S07).

Some genes of note are missing from single strains including; STY0688 – *dacA,* D-alanine carboxypeptidase from strain S06, SBG2553 – *ygbF,* encoding a conserved hypothetical protein from strain S07, and STY4668, a conserved hypothetical protein gene associated with SPI7 from strain S21.

Table 4-1 **Percentage similarity of the 32 test strains to sequenced DT104 and the number of missing or addition genes. Data obtained from comparative genomic microarray and processed with Microsoft Excel**

Sample	No. genes missing	% to DT104	No. Additional genes	% of DT104	
S01	17	99.57	28	0.71	Human Resistant Strains
S02	22	99.44	22	0.56	
S03	18	99.55	22	0.56	
S04	26	99.34	3	0.08	
S05	21	99.47	13	0.33	
S06	22	99.44	11	0.28	
S07	25	99.37	6	0.15	
S21	28	99.29	4	0.10	
S22	27	99.32	9	0.23	
S24	25	99.37	22	0.56	
D01	53	98.66	6	0.15	Animal Resistant Strains (1)
D02	69	98.26	2	0.05	
D03	69	98.26	3	0.08	
D07	39	99.02	2	0.05	
D08	43	98.91	4	0.10	
D09	50	98.74	8	0.20	
D04	57	98.56	40	1.01	Animal Resistant Strains (2)
D05	57	98.56	95	2.40	
D10	63	98.41	56	1.41	
D11	55	98.61	62	1.56	
D12	52	98.69	55	1.39	
S13	93	97.65	2	0.05	Sensitive Strains (1)
S16	59	98.51	7	0.18	
S19	79	98.01	3	0.08	
S27	89	97.75	3	0.08	
S28	80	97.98	5	0.13	
S29	76	98.08	8	0.20	
S30	86	97.83	6	0.15	
A01	140	96.47	27	0.68	Sensitive Strains (2)
A02	268	93.24	45	1.14	
S15	227	94.27	78	1.97	
D06	160	95.96	50	1.26	
LT2	137	96.54	27	0.68	Control Strains
SL1344	134	96.62	257	6.48	
PT4	352	91.12	198	5.00	
S.Gal	392	90.11	208	5.25	
S.Bong	1019	74.29	483	12.19	
CT18	504	87.28	812	20.49	

There are also some additional genes in these strains when compared to the sequenced DT104. S01 has 14 genes associated with a *S. bongori* phage and S02 has 18 genes associated with the SL1344 plasmid pSLP2. Strains S05 and S06 have the HCM1.223 - *strA* and HCM1. 224 - *strB* streptomycin resistance genes and the HCM1.165c - *dhfR* – dihydrofolate reductase and HCM1.222 - dihydropteroate synthase genes associated with the *S.* Typhi CT18 HCM1 plasmid. S02 also has the same *strA* gene whilst S07 has a beta-lactamase gene, ^{bla}*TEM,* also identified as coming from the same source. Strain S24 has an additional glutaredoxin HCM2.0023c. The effect of the resistance genes in strains will be considered below.

The second group contains 6 animal resistant strains (D01 - D03, D07 - D09) with between 98.26% to 99.02% similarity to the sequenced DT104. These strains have a number of commonly missing genes including those related to the Fels-1 prophage (STM0899, STM0900 and STM0918) as well as genes associated with the Gifsy-2 prophage (SEN0911 and SEN0912). Two of the strains, D02 and D03, lack 14 genes between STM2584 and STM2636, which are classified as being from the Gifsy-1 phage. They also lack genes associated with an ADP ribosyltransferase 'pertussis like' region STY1041, STY1042, STY1362 and STY1365. All the strains show the absence of STY1364 which may also be related to this toxin region.

Individual strains also show absences of notable genes. Strain D02 lacks AF071555_tetR and its homolog AF071555_tetR_DT104 which code for tetracycline resistance. D02 also lacks the bla^{PSE} as well which would confer beta lactam resistance. Strain D09 lacks STY1825 – *gapA,* a glyceraldehyde 3-phosphate dehydrogenase A and STY3466 – *rbfA*, a ribosome-binding factor. Strain D03 lacks

STY1913 – *hyaA*, hydrogenase-1 small subunit and STY3961 – *ccmC2*, heme exporter protein C2.

There are few additional genes identified for these strains. Three strains, D01 - D03 contain, STY0577 – *fdrA*. Strain D01 contains additional genes from the *S*. Typhi HCM1 plasmid, HCM1.165c – *dhfR* a putative dihydrofolate reductase, HCM1.217 – a hypothetical protein, HCM1.222 – a dihydropteroate synthase, HCM1.223 – *strA*, streptomycin phosphotransferase and HCM1.224 – *strB*, streptomycin phosphoreductase.

The third group, Animal Resistant (2), was separated from the first animal resistant group because of the higher number of additional genes, between 40 and 95, when compared to the sequenced DT104. These strains are missing genes related to some mobile genetic elements, from bacteriophage sources such as Fels1 and Gifsy2 and other prophage regions. Multiple strains are also lacking genes from the SPI7 region.

Other genes of note missing from individual strains include: STY1682 – *sodC*, a superoxide dismutase that was missing from strains D10, D11 and D12 and STY1913 – *hyaA*, hydrogenase 1 small subunit and STY1929 – a putative Na^+/H^+ exchanger missing from strains D10 and D11. There was also variation within the operon for cobalamin biosynthesis, D10 lacks STY2232 - *cbiH*, precorrin-3 C17-methyltransferase and STY2223 - *cbiO*, a putative cobalt transport ATP-binding protein. D12 lacks STY2228 - *cbiL*, precorrin-2 c20 methyltransferase. Strain D10 also lacks the genes STY2243 - *pduA*, putative propanediol utilization protein, STY2247 - *pduE*, diol dehydratase small subunit, STY2290 - *gnd*, 6-

phosphogluconate dehydrogenase and STY2425 - *galS*, galactose ultrainduction factor.

As noted these strains have more additional genes when compared to the other two resistant groups. The majority of these originate from the mobile genetic elements present in Enterobacteriaceae including plasmids, bacteriophages and pathogenicity islands. Some of these genes are related to antibiotic resistances from the *S*. Typhi plasmids HCM1 and HCM2.

The fourth group contains the majority of the human derived sensitive strains; S13, S16, S19, S27 - S30, from both the pre-epidemic and epidemic period. These have few additional, genes when compared to DT104, but do have between 59 and 93 genes missing. Some of these genes are missing for all strains of which 38 represent genes from the *Salmonella* Genomic Island 1 (SGI1) – AF071555 genes, AF261825 genes, SDT3826 – SDT3864, which confers antibiotic resistance. Two common genes, HCM1.162c and HCM1.167, represent a transposase and integrase respectively. They also lack genes SEN0911 and SEN0912 associated with Gifsy-2.

All strains, with the exception of S16, lack SBG0906 a hypothetical phage protein. The gene STY1364, which has been associated with a pertussis like toxin, was missing from strains S13, S27 - S30. Other notable gene absences are STY3961 – *ccmC2*, a heme exporter protein C2 and STY3964 – *ccmF2*, a cytochrome C type biogenesis protein F2 from strains S13 and S28. Strains S13 and S19 are missing the genes STY3006 – *sipD*, a SPI1 effector protein and STY3725 – *thiG*, a thiamine biosynthesis protein. S13 also lacks two other genes associated with SPI6: STY3007 – *sipC*, encoding an effector protein and *spaO*, encoding a surface presentation of

antigens protein, and also lacks STY4794 – *treB*, a PTS system trehalose specific IIBC compound.

Some of the missing genes have postulated roles in metabolism. Strains S28 and S29 are lacking the gene STY2698 – *eutG*, a putative alchohol dehydrogenase involved in ethanolamine utilisation. S27 shows the absence of the D-ribose operon, STY3889 – STY3895, including the genes *rbsR*, operon repressor, *rbsK*, ribokinase, *rbsB*, D-ribose binding periplasmic protein and *rbsC*, a high affinity ribose transport protein.

The fifth group consists of four strains which show between 93.24 and 96.47% identity to the sequenced DT104. These show a high number of between 140 – 268 genes missing, and between 27 – 78 additional genes. Common absences include SGI1, HCM1.162c and HCM1.167, genes from phage ST104 (SDT0325, SDT0340, SDT0343, SDT0356) and prophage 3 (SDT1840 - SDT1843, SDT1865, SDT1868, SDT1870).

Strains A01, A02 and S15 lack prophage 10 (STY1040-1046) as well as genes with an association with a pertussis like toxin (SBG0912, SDT2668, STY1362 – 1365). Absences shared by two strains include the LT2 plasmid (PLST001-PLSTtrb), *Salmonella* pathogenicity island 6 (STY0286-STY0328) and genes from pathogenicity islands 7 and 10, for strains S15 and A02, prophage 3 for strains A02, D06 and S15 as well as the phage ST104 from strains A01 and A02.

These strains have a number of additional genes when compared to DT104. Common to all four strains was a sequence of 11 genes between STY0566 and

STY0577 that have been associated with the utilisation of allantoin as a nitrogen source and to glyoxylate metabolism. This will be considered further below.

A02 and S15 have a significant number of genes for SPI6 from the *S.* Gallinarum gene set, SG0263 – SG0300. S15 shows the presence of the enterobacteria phage P2, SBG2807 – BG2845, genes associated with the *S.* Typhi SPI6 region, STY0344 – STY0350 and genes from a *S.* Typhimurium LT2 phage region, SL2638 – SL2643. D06 shows the presence of 4 genes from the HCM1 plasmid, HCM1.108, HCM1.270, HCM1.271 and HCM1.281c.

In total these strains are missing 324 distinct genes when compared to the sequenced DT104 strain, of these genes 261 have known association with phages, pathogenicity islands, plasmids or the SGI1 region. Of the 129 additional genes 93 have the same type of association with mobile or mobilizable elements.

It should be noted that all the strains in the study lack a series of genes related to prophage regions associated with *S.* Enteritidis; SEN1379, SEN1388, SEN1393 – SEN1394 from the prophage SE14 and 6 genes between SEN1936 and SEN1959 associated with SE20. These genes were shown to be present in the sequenced DT104, but have previously only been reported in *S.* Enteritidis

Discussion

The results from the microarray analysis show that genetically the strains generally group along resistance rather than any epidemic lines. Resistant strains from the epidemic period group together with those resistant strains from the post epidemic period. Within the resistance group there was some differentiation between human

and animal isolates. Sensitive strains from the pre-epidemic period group with those from the post-epidemic period. The major factor which differentiates between strains was the presence or absence of the SGI1 chromosomal island which confers antibiotic resistance. However some other patterns of difference are observable.

The majority of variation that exists among the strains lies with genes associated with mobile elements. These differences predominantly involve plasmids, bacteriophages, prophages or pathogenicity islands. The SGI1 itself, although chromosomally located, has been shown to be mobilizable (Doublet et al., 2005a). This suggests that the variation within DT104 is probably predominantly driven by horizontal gene transfer. Some regions such as that for allantoin and glyoxylate utilisation and a 'pertussis like' toxin region are interesting. The loss of some genomic genes in individual strains, especially those with a possible metabolic cost, may be reflecting the selection pressures of cold storage.

4.3 Antibiotic resistance

Introduction

In the previous section, variability was shown with respect to antibiotic resistance genes. It is well established that the resistance genes in the classic penta resistant epidemic DT104 types were related to the presence of a specific genomic island named SGI-1 that encodes resistance and virulence determinants (Boyd et al., 2001) and that some variants have also been described (Doublet et al., 2005a). Here, the aim was to describe in more detail the gene content related to SGI-1 present and absent in the test panel.

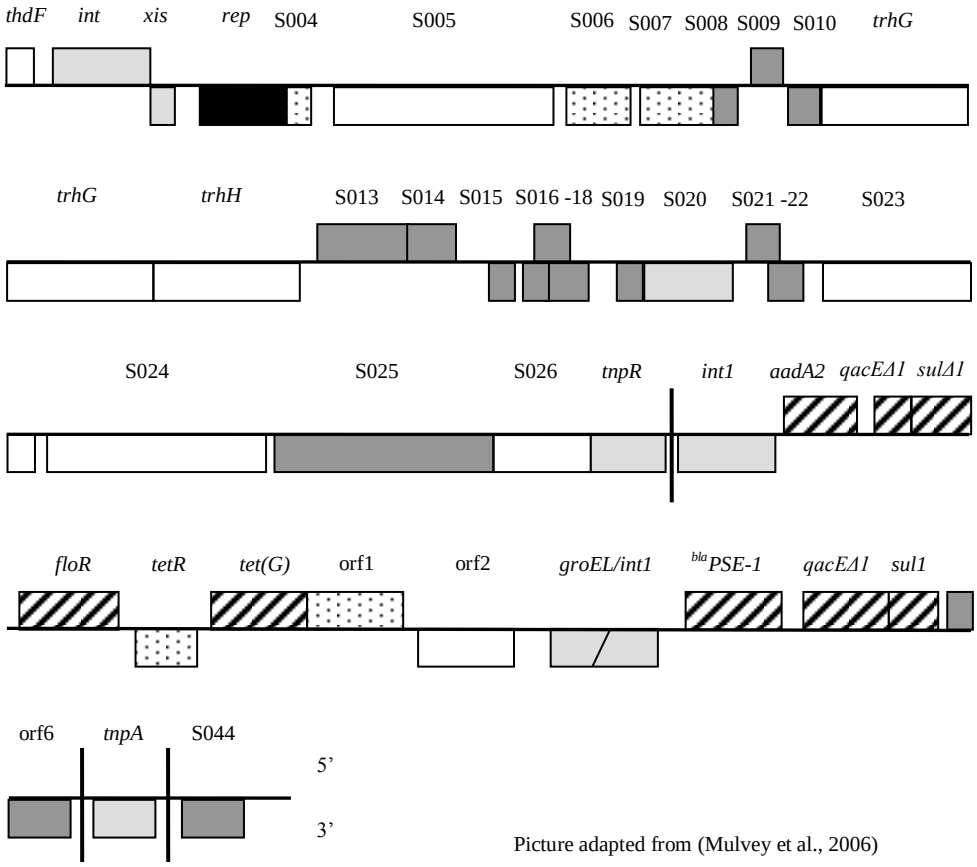
4.3.1 Variation within SGI-1

The strains with the traditional ACSSuSpT profile, S01 - S03, S05, S07, D01, D03 – D05, D07 – D11, show the presence of all the genes from the SGI1 region mirroring the sequenced gene cassettes from DT104 (Figure 4-2). These consist of the genes; *aadA2, qacEΔ1, sulΔ1, floR, tetR, tet(G),* ^{bla}*PSE* and *sul1*. The *aadA2* gene codes for an aminoglycoside adenyltransferase, which mediates resistance to the aminoglycoside streptomycin and the aminocyclitol spectinomycin. Streptomycin adenylyltransferase catalyses the transfer of adenyl group from ATP into the 3" and 6' position of the streptomycin molecule, inactivating the enzyme (Jana and Deb, 2005). Although the presence of copies of the *qacEΔ1* gene would suggest an increased resistance to quaternary ammonium compounds, no such increased resistance has been seen in *E. coli* harbouring this gene (Kucken et al., 2000).

The *floR* gene codes for a protein with shows homology to the efflux pump encoded by *cmlA* gene of *Pseudomonas aeruginosa*, and gives resistance to both chloramphenicol and the fluorinated structural analogue of choramphenicol, florfenicol (Cloeckaert and Schwarz, 2001; Doublet et al., 2005b). The tetracycline resistance gene *tet(G)(*or *tetA*) codes for an efflux pump, whilst *tetR* codes for a repressor which controls its expression (McMurry et al., 1980). The ^{bla}*PSE-1* gene codes for a beta-lactamase enzyme that was first characterized in *Pseudomonas aeruginosa* as hydrolysing carbenicillin and a number of other B-lactam antibiotics (Matthew and Sykes, 1977). The *sul1* gene codes for a dihydropteroate synthase which shows a pronounced insensitivity to sulphonamide inhibition (Skold, 2001). In combination these genes confer the resistance profile ACSSuSpT.

Variations can be distinguished in the genes making up the SGI1 region. While gene AF071555_*aad*A2 (for streptomycin/spectinomycin adenyltransferase) was absent from DT104 strains S15 and A02 as well as the control strains *S.* Gallinarum, *S.* Typhi CT18, *S. bongori*, it was present in the rest of the strains. The homologue for AF071555_*aad*A2, SDT3850, was only present for those DT104 strains that have antibiotic resistant profiles. Variant resistance profiles were also identified by the microarray data. Strains D02, D12, S21, S22 and S24 which have the profile SSu showed only the presence of the resistance genes *aadA2*, *qacEΔ1, sulΔ1* and *sul1*, conferring resistance to streptomycin and sulphonamides. The S04 strain showed similar variation with *aadA2, qacEΔ1, sulΔ1, floR* and *sul1*. These genetic variations were mirrored by variation in the resistance profiles confirmed by the antibacterial MIC.

Figure 4-2 Pictorial representation of the *Salmonella* Genomic Island 1. Genes for antibiotic resistance are represented by the hashed boxes



Picture adapted from (Mulvey et al., 2006)

4.3.2 Other antibiotic resistances

Other genes, *dhfR*, *sul-2*, *strA*, *strB* and ^{bla}*TEM*, which confer antibiotic resistances, were detected by the microarray results in strains S05, S06 and D01. These were identified from genes on the array that were from the *S.* Typhi CT18 HCM1 plasmid. This is an *inc*HI1 antimicrobial resistance plasmid first isolated from cases of multidrug resistant typhoid fever in Vietnam in 1993 (Wain et al., 2003), that contains antibiotic resistances on gene cassettes as well as genes for mercury resistance (Phan et al., 2009). Hybridisation with these specific gene probes does not necessarily imply that the presence of the *S.* Typhi CT18 HCM1 plasmid but rather that similar resistance genes were present.

The *dhfR* gene encodes a dihydrofolate reductase and as such could provide the resistance to trimethoprim present in strain S05. Trimethoprim acts as a dihydrofolate reductase inhibitor and therefore interferes with purine and thymidylate biosynthesis and ultimately cell division and proliferation. *Sul-2* encodes a dihydropteroate synthase providing resistance to sulphonamides. Two genes *strA* and *strB* code for enzymes for resistance to aminoglycoside antibiotics such as streptomycin. Also detected was a ^{bla}TEM coding for a beta-lactamase and conferring resistance to beta-lactam antibiotic penicillins including ampicillin. Both ^{bla}*TEM* and *sul-2* are usually located on plasmids (Skold, 2001).

Discussion

The CGH microarray data confirms the presence of the genes that make up the classical SGI1 in the 14 strains with the penta-resistant profile. Those strains with variant profiles that showed fewer resistances also showed variation with the loss of some genes from this region. One limitation of the comparative genomics technique is that novel additions and re-arrangements within the SGI1 would not be detected. To interrogate possible changes within this region more comprehensively a sequencing based technique would have to be used. However from the microarray data it is reasonable to assign known variant structures to some strains.

Since the composition of the classical *Salmonella* Typhimurium DT104 SGI1 has been described (Boyd et al., 2001) other examples with different resistance profiles have been identified. Variants SGI-A to SGI-O have been described in Typhimurium and also in *Proteus mirabilis,* and variant genomic island with a different lineage, termed SGI2 has also been described (Boyd, 2008; Levings, 2008;

Mulvey et al., 2006). The genetic results and the resistance profiles ascertained in the study suggest that strains D02, S21, S22 and S24 have the variant SGI1-C (Mulvey et al., 2006).

The presence of the resistance genes ^{bla}*TEM*, *strA*, *strB* and *sul2* in strains S05, S06 and D01 is suggestive of a unique resistance determinant region. These genes have recently been associated with a clonal group of *S.* Typhimurium and its monophasic variant *Salmonella enterica* subsp. *enterica* serovar 4,[5],12:i:- that emerged during the 2000's. These strains display the resistance profile ASSuT and the resistance genes have been shown to occur in a separate and distinct chromosomally encoded region when compared to SGI1 (Lucarelli et al., 2010). The presence of a subset of these HCM1 genes, but not the whole plasmid, suggests that these strains may contain this newly identified region or another plasmid or plasmids with these genes.

4.4 Genomic differences

Introduction

The microarray data has confirmed the resistance profiles and also uncovered variation within genes related to antibiotic resistance. The acquisition of antibiotic resistance by DT104 is thought to have coincided with the emergence of the epidemic. However other differences have been uncovered by the microarray study that could also affect the relative fitness of the strains and therefore may have an impact on their epidemicity.

4.4.1 Allantoin and Glyoxylate

S. Typhimurium LT2 contains a region of 11 genes, STY0566-STY0577, which was absent from the sequenced DT104. The list of these genes and their predicted functions are listed in Table 4-2.

Table 4-2 **List of genes STY0566-STY0577 present in *S.* Typhimurium LT2 but absent from the sequenced DT104 genome. These genes are involved in glyoxylate and allantoin utilisation**

Gene I.D.	Function
STY0566	*hyi*, hydroxypyruvate isomerase, glyoxylate induced protein
STY0567	*glxR*, 2-hydroxy-2-oxopropionate reductase
STY0568	Probable metabolite transport protein
STY0569	*ybbW*, Putative allantoin permease
STY0571	*allB*, putative allantoinase
STY0572	Putative permease
STY0573	*glxK*, glycerate kinase
STY0574	*Conserved hypothetical protein*
STY0575	*allC*, allantoate amidohydrolase
STY0576	*allD*, ureidoglycolate dehydrogenase
STY0577	*fdrA*, involved in protein transport

STY0566 – STY0577, are absent from the majority of resistant and sensitive test strains but are found in human sensitive strain S15 and animal sensitive strains A01, A02 and D06, as well the control strains LT2, SL1344, PT4, *S.* Gallinarum and CT18. This region has been previously identified as being present in some strains of DT104 that are sensitive, and it has been hypothesised that the loss of this region pre-dates the acquisition of SGI1 in resistant strains (Matiasovicova et al., 2007).

4.4.2 Ribose

Of all the strains tested by microarray only human sensitive strain S27 shows the absence of the D-ribose operon; STY3889, STY3891, STY3892, STY3894 – STY3895. The gene identifications are listed in Table 4-3.

Table 4-3 **List of genes STY3889-STY3895 present in *S.* Typhimurium strains including DT104.. These genes are associated with the ribose operon. They are missing from human sensitive strain S27**

Gene I.D.	Function
STY3889	*yieO*,putative transmembraneefflux protein
STY3891	*rbsR,* ribose operon repressor
STY3892	*rbsK*, ribokinase
STY3894	*rbsB*, D-ribose-binding periplasmic protein
STY3895	*rbsC*, high affinity ribose transport protein

4.4.3 'Pertussis like' region

Another area of interest that arises from the study is the presence or absence of a series of genes coding for a predicted pertussis like toxin similar to the one from *Bordetella pertussis*. This region has been characterised in DT104 as containing an ADP-ribosyltransferase toxin homologue (*artA, artB*) that is flanked by prophage like sequences suggesting acquisition by phage mediated recombination (Saitoh, 2005).

Figure 4-3 shows variation with genes associated with the pertussis-like toxin region. All the animal strains, two of the sensitive human strains S13 and S15, as well as the pre-epidemic sensitive strains S27-S30 lack the gene STY1364 which encodes *artB*, one part of the *artA/B* toxin gene pair. Sensitive strains A01, A02, D02, D03 and S15 also lack genes surrounding this gene, STY1362 and STY1365.

Other gene homologues of the 7 open reading frames identified by Saitoh et al in the DT104 ADP-ribosyltransferase region show the same pattern of presence. Orf5 - STY1040, a holin, and orf6 - STY1041, a prophage membrane protein are missing from A01, A02, D02, D03 and S15. The same pattern exists for gene SBG0912 which shows a 94% identity to the pertussis gene *artA/B* and its flanking region.

Figure 4-3 Presence of genes STY1362, STY1364 and STY1365 associated with a 'pertussis like' toxin in the 32 test strains from the comparative genomic microarray

	Strain																														
Gene ID	A01	A02	D01	D02	D03	D04	D05	D06	D07	D08	D09	D10	D11	D12	S01	S02	S03	S04	S05	S06	S07	S13	S15	S16	S19	S21	S22	S24	S27	S28	S29	S30
STY1362	-	-	+	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+
STY1364	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	-	-	+	+	+	+	+	-	-	-	-
STY1365	-	-	+	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+

4.4.4 PCR for *artA/B* presence in strains

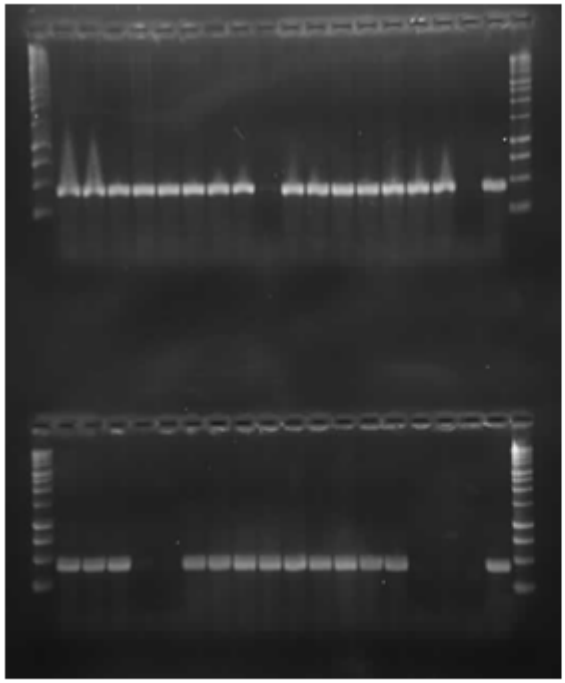
Using the *artA/B* toxin region identified in DT104 (Saitoh, 2005) and the available sequence of DT104 (Dougan), PCR primers were designed to test the presence *art*A/B genes. A PCR was carried out for all strains in the study.

Figure 4-4 shows the results for the *artA/B* bridging region. The absence of the region was confirmed in strains S15, D02, D03, A01, A02 and the negative control strain LT2. This supports the results from the microarray (Figure 4-3 above) which shows that these strains lack all three of the genes; STY1362, 1364 and 1365.

Sequencing of the *art*A/B region was carried out on the 10 strains; S01, S04, S05, S13, S21, S24, S27, D01, D10 and the sequenced DT104. These were chosen because they showed either total or partial presence of the region from the CGH

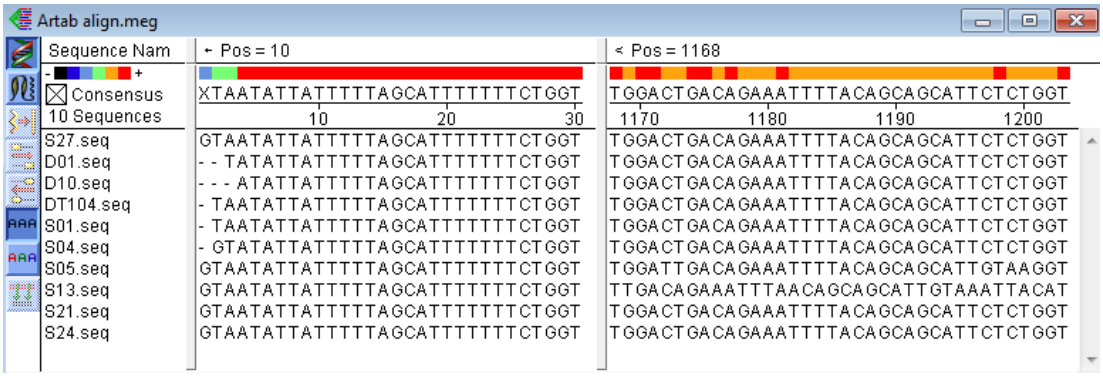
microarray results (Figure 4-3 above). The completed sequences were processed with DNASTAR Lasergene SeqMan and MegAlign. Figure 4-5 shows the alignment of the strains from a screen grab from the MegAlign software.

Figure 4-4 **Electrophoresis 1% agarose gel with ethidium bromide staining. PCR product for *artA/B* bridging regions for all 32 test strains run on gel for 1hr with a 1kb ladder. Expected product size 1340b**



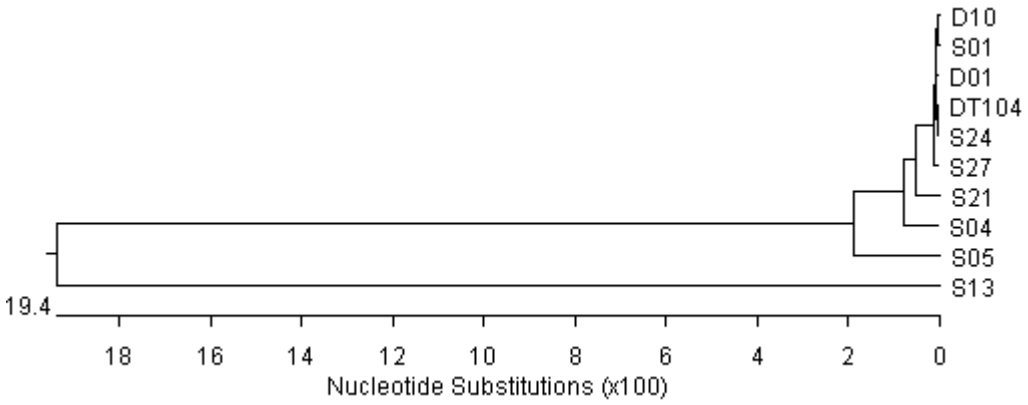
Row no.	Sample	Row no.	Sample
1	1kb ladder	21	1kb ladder
2	S01	22	S29
3	S02	23	S30
4	S03	24	D01
5	S04	25	D02
6	S05	26	D03
7	S06	27	D04
8	S07	28	D05
9	S13	29	D06
10	S15	30	D07
11	S16	31	D08
12	S19	32	D09
13	S21	33	D10
14	S22	34	D11
15	S24	35	D12
16	S27	36	A01
17	S28	37	A02
18	LT2	38	LT2
19	DT104	39	DT104
20	1kb ladder	40	1kb ladder

Figure 4-5 **Alignment of a subset of 10 of the test strains using their obtained sequences in the MegAlign software**



The ClustalW phylogenetic tree grouping of the strains is shown in Figure 4-6. The strain that diverges most from the others was human sensitive strain S13 from the later period. The next most divergent are human resistant strains S04 and S05 from the epidemic period. S21 a later resistant strain, also shows some divergence. The rest of the strains group closer together. The sequencing data showed that there was variation within the *artA/B* region. The greatest variation was with S13, which has shown variance from the CGH microarray data. Other strains in which variation was predicted, D01, D10, S27, in fact showed little difference from strains that showed the presence of the whole region.

Figure 4-6 **ClustalW phylogenetic tree of sequenced *artA/B* region from the MegAlign software, showing variation in the number of nucleotide substitutions and the relationship between the strains variation.**



4.4.5 *pdu/cob* regulon

Animal strain D10 shows the absence of some genes from a region associated with the biosynthesis of adenosyl-cobalamin (vitamin B_{12}) and with the catabolism of 1,2-propanediol. The *cob* operon for B_{12} synthesis consists of 25 genes and is split into three groups, each encoding enzymes for a distinct part of the pathway (Roth et al., 1993). The missing genes c*biO*, a putative cobalt transport ATP binding protein and *cbiH*, precorrin-3 C17-methyltransferase, come from group I that bring about the synthesis of cobinamide. The synthesis of only a few enzymes in *E. coli* are B_{12} dependent including ethanolamine deaminase and propanediol dehydratase (Roof and Roth, 1989).

Adjacent to the *cob* operon is the *pdu* operon, which controls the breakdown of 1,2-propanediol to propionaldehyde. This can be used as hydrogen sink when converted to propanol and excreted or can be used as a carbon source converting to pyruvate via propionyl-coenzyme A. These two operons are induced by propanediol, via the activator *pocR*, and are under positive control of the global regulatory proteins Crp and ArcA (Chen et al., 1994). Strain D10 was missing the genes *pduA*, propanediol utilization protein and *pduE*, a diol dehydratase small subunit.

4.4.6 Mobile elements

4.4.6.1 Plasmids

The 93 kb virulence plasmid from LT2, pSLT, is included on the CGH microarray and was shown to be present in all strains except animal sensitive strain A02 and human sensitive strain S15. Plasmid pSLT and the virulence plasmids from other

serovars including SL1344 have been shown to be self-transmissible, although in the latter case at a much lower frequency (Ahmer et al., 1999; Garcia-Quintanilla and Casadesus, 2011). The lack of this region in these strains is interesting, but has not been confirmed by plasmid extraction, and this would be a target for further work.

In addition to pSLT human resistant strain S02 possesses 18 genes of the 103 genes representing the *S.* Typhimurium SL1344 plasmid SLP2. This may be indicative of a distinctive plasmid or a truncated/chimeric derivative in this strain. Other plasmids in *S.* Typhimurium serovars have been shown to carry integron-borne gene cassettes and both virulence determinants and antibiotic drug resistance genes (Guerra et al., 2002). The presence of unique plasmids may therefore have implications for epidemicity.

4.4.6.2 Fels-1 phage

Figure 4-7 shows the presence and absence of genes associated with the Fels-1 phage (STM0893-0929). The strains with animal origin D01-D12 show the general absence of these genes. S13, a sensitive human strain, also shows a greater variation in this region when compared to the other human strains.

Figure 4-7 Presence or absence of genes STM0893-STM0929 associated with the Fels-1 phage in the 32 test strains from the comparative genomic microarray

Gene ID	A01	A02	D01	D02	D03	D04	D05	D06	D07	D08	D09	D10	D11	D12	S01	S02	S03	S04	S05	S06	S07	S13	S15	S16	S19	S21	S22	S24	S27	S28	S29	S30	DT104	LT2
STM0899	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	-	-	+	-	+	+	+	+	+	+	-	+	+
STM0900	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	-	+	+	-	+	+	+	-	+	+	-	+	+
STM0906	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
STM0918	+	+	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+	+

4.4.6.3 Gifsy-1 phage

Animal resistant strains D02 and D03 have been shown to be negative for Gifsy-1. The lambdoid prophage Gifsy-1 from *S*. Typhimurium has been shown to effect virulence in strains that lack the Gifsy-2 prophage region. Curing *S*. Typhimurium of Gifsy-1 and Gifsy-2 reduces the virulence of the bacterium in mouse models (Figueroa-Bossi and Bossi, 1999).

Discussion

The genetic data from the microarray has shown some differences between strains that may have implications for both strain fitness and virulence. The presence of genes involved in allantoin and glyoxylate utilisation may have an important impact on the bacterial metabolism and their loss in the 'epidemic' resistant strains may have significant implications for epidemicity. The loss of this region has been reported to have pre-dated the acquisition of the antibiotic resistant region SGI1 (Matiasovicova et al., 2007). It may therefore play an important role in the evolution of the epidemic strain. The loss of the ribose operon from one strain and the *pdu/cob* operon from another may have significance.

The absence of the region associated with a 'pertussis' like toxin from some sensitive strains was also marked. This toxin region has been reported to be inactive in DT104 because of at least one frame shift (Hermans, 2005a; Parkhill et al., 2001). However a recent study has shown that the genes in this region can be transcribed in DT104 and produce an ADP ribosyltransferase (Uchida, 2009). Therefore the presence of this region in DT104 may have a more important role in virulence and epidemicity than first thought.

The comparative genomic hybridisation array has revealed genetic differences between the strains. The drawbacks of the technique have also become apparent; it is unable to identify novel genes, it would not identify mutations within genes or rearrangements within regions. The results involving mobile elements suggest that cross hybridisation with common phages may also confuse the overall picture.

Overarching these problems is the fact that the presence of a gene does not automatically mean it is actively transcribed. For that reason, and to confirm the genetic results, a comprehensive study, using the Biolog phenotyping microarray, was carried out.

4.5 Optical Mapping

Introduction

Another question arising from the studies above and not answerable by CGH alone is the possibility of genomic rearrangements. Whilst CGH approaches can give sound data on the presence or absence of genes, it is not able to provide any relational or spatial data. As described above, it has been possible to postulate that certain gene combinations are likely to reflect certain gene relational structures such as in the SGI1 and its variants. In addition the presence of phage related integron and excision genes, but not the whole phage related to them on the array, suggests that novel phages may be present. However, without definitive data from sequence approaches these can only be speculative deductions and therefore prone to drawing inaccurate conclusions. During the course of this study, the opportunity to utilise a new system of Optical Genetic Mapping to establish the physical maps of the various genomes became available. Here, a subset of 9 of the DT104 strains, along with field isolates from the

serovars *S.* Enteritidis (1 strain), *S.* Newport (7 strains) and *S.* Dublin (3 strains), were investigated with a novel technology, optical genetic mapping. The data for all three serotypes have now been published (Saunders et al., 2010). The approach generates a physical map of the whole bacterial chromosome for the comparison of strains of *Salmonella* to each other and the available sequenced strains.

This method has the capacity to combine greater resolution, positional information and the identification of novel insertions to an extent that is lacking from techniques such as PFGE. In brief, specific restriction endonucleases cut the DNA of a bacterium that has been stretched out and immobilised on a derivatized glass slide. The restriction fragments are fluorescently labelled *in situ*, and then are visualised and photographed. The mass of each fragment is determined by the intensity and partial genome maps are developed. These are assembled by overlapping segments into a genome optical map using alignment software. Optical maps for the 20 *Salmonella* serovars were prepared by OpGen, Gaithersburg, MD, USA following the method presented in Zhou *et al*. (2004) (Zhou, 2004)

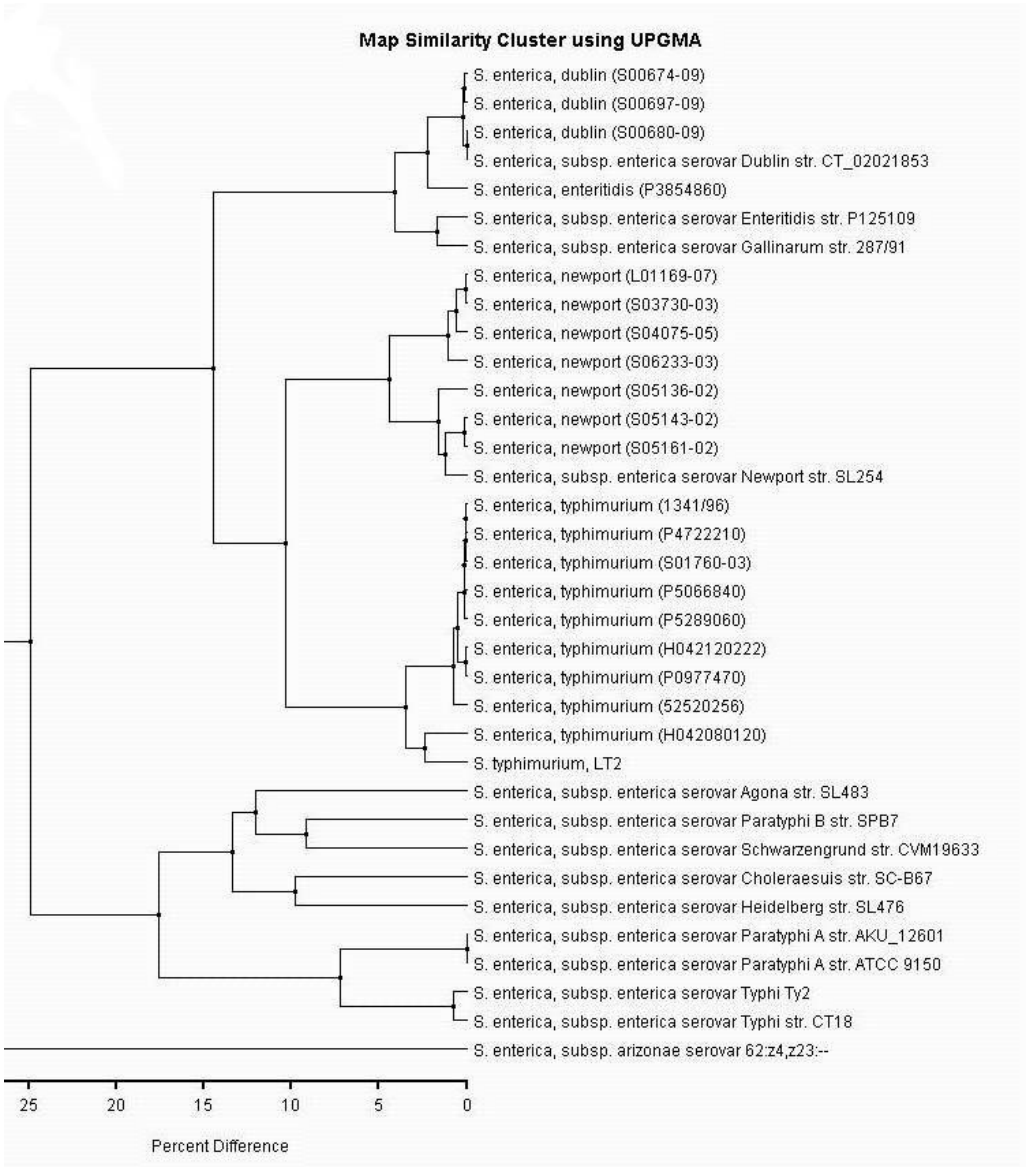
4.5.1 Genome alignment relationships.

The OpGen MapSolver software was used to create an unweighted pair group method with arithmetic mean (UPGMA) to create a phylogenetic tree of the optical maps from the available *in-silco Salmonella* sequences and those generated from the *S.* Typhimurium and other test strains (Figure 4-8). The clustering correctly grouped the test strains with their respective 'control' sequenced strains. The 9 *S.* Typhimurium test strains showed little variation with a 0.8% difference within the group. The exception to this was strain S15 which grouped closer to the LT2 control strain. The sequenced *S.* Dublin and 3 *S.* Dublin test strains showed the least variation,

with 0.3% difference within the group. The 7 *S.* Newport test strains generally clustered into two groups which will be briefly discussed below. The *S.* Enteritidis test phage type 11 strain P3854860 clustered closer to the *S.* Dublin strains than to the control *S.* Enteritidis phage type 4 strain P125109, which was closer to the sequenced Gallinarum strain with a 1.8% difference.

The *in-silico* optical maps of the sequenced *S.* Typhimurium LT2 and multidrug resistant DT104 were compared (Figure 4-9). The regions identified as mobile elements and their genomic coordinates for LT2 by Hermans et al (Hermans, 2005a) and for DT104 by Cooke et al (Cooke et al., 2007), are highlighted. The variation between the strains was predominantly in these regions associated with mobile elements, notably prophages. *S.* Typhimurium LT2 contains the regions of the bacteriophages Fels-1 and Fels-2 that are absent from DT104. Fels-1 has been shown to contain the genes *sodCIII* (superoxide dismutase), *nanH* (neuraminidase) and *grvA*. DT104 contains prophage 1 (ST104), prophage 3, prophage 4 and the *Salmonella* Genomic Island 1 (SGI1) that are absent from the LT2 strain. This later region contains the genes that confer antibiotic resistance to DT104 strains (Boyd et al., 2001). Both strains show the presence of Gifsy-1 and Gifsy-2, two lambda-like phages that have been associated with virulence (Brüssow, 2004).

Figure 4-8 **Unweighted pair group method with arithmetic mean (UPGMA) phylogenetic tree of the optical maps from the available *in-silco Salmonella* sequences and those generated from the *S.* Typhimurium and other test strains. Showing percentage difference between strains**

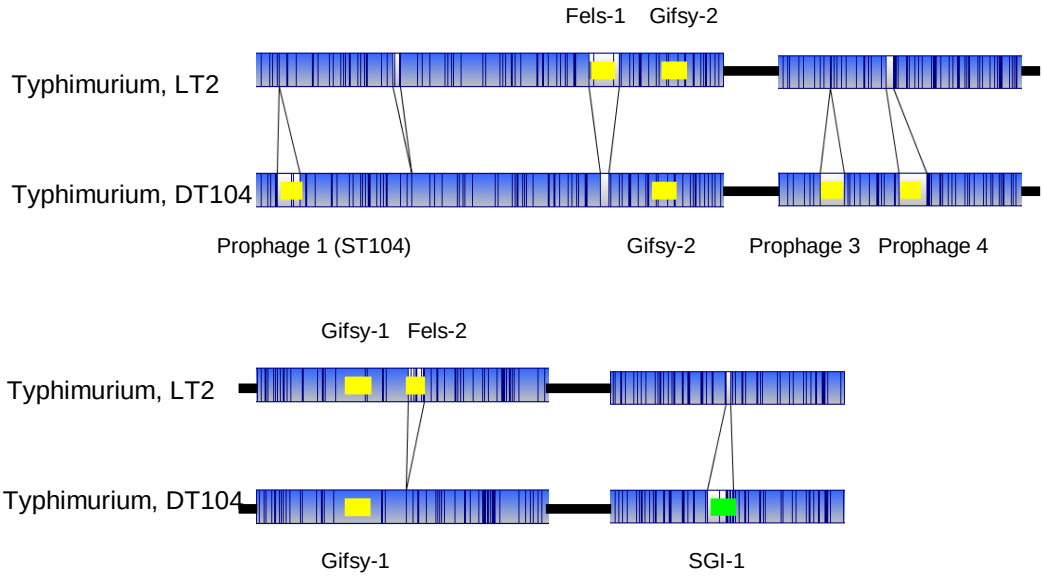


4.5.2 Variations within the *S*. Typhimurium test strains.

Nine of the test strains were analysed with this technique, S01, S04, S06, S15, S16, S221, S27, D01 and D07. The upper part of Figure 4-10 shows regions of variation between the *S*. Typhimurium strains S21 and the sequenced DT104. A 40kb insert downstream from prophage 5 was identified, that was absent from the sequenced strain. This insert was located upstream of genes related to purine metabolism (*purM* and *purN*) and *ppk* (polyphosphate kinase) and downstream of *guaB* (inositol-5-monophosphate dehydrogenase).

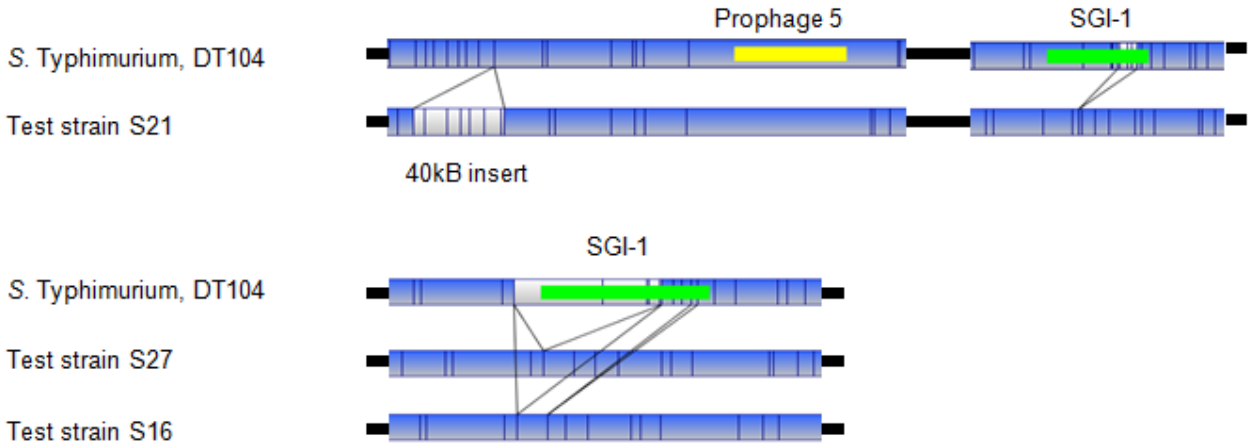
Strains S21 (Figure 4-10) and S04 (not shown) are antibiotic resistant but have different profiles from the penta-resistant profile of the sequenced strain. These differences in resistance profile were visible by optical mapping, with partial deletions in the region of SGI-1 in the test strains, between bases 4,115,969 and 4,124,741 for strain S21 and between bases 4,112,582 and 4,122,042 for strain S04. This region encompasses the second 'resistance cassette' of the genomic island including, for S21, the genes *tet(G)*, *groE1/int1*, ^{bla}*pse-1*, *qacEΔ1, sul1* and *tnpA*, and for S04 the genes *floR, tetR, tet(G), groE1/int1*, ^{bla}*pse-1* and *qacEΔ1* (Boyd et al., 2001).

Figure 4-9 **Comparison of the optical maps of control LT2 and DT104 strains. Yellow blocks represent the identified phage and prophage regions of LT2 and DT104. The green region represents the *Salmonella* Genomic Island 1. The genomic location of the regions is listed in the table below**



Identifier	Genome position, base pair number	
DT104 (Cooke et al, 2007)		
Prophage 1 (ST104)	365,588	406,646
Prophage 2 (Gifsy-2)	1,079,201	1,124,674
Prophage 3	1,954,176	1,995,393
Prophage 4	2,110,069	2,149,002
Prophage 5 (Gifsy-1)	2,797,168	2,845,750
SGI1	4,083,657	4,129,583
LT2 (Hermans et al, 2006)		
Prophage Gifsy-1	2,726,772	2,777,314
Prophage Gifsy-2	1,096,831	1,144,131
Prophage Fels-1	962,001	1,005,353
Prophage Fels-2	2,844,427	2,878,156

Figure 4-10 **Comparison of the optical maps of the control DT104 strain and 3 DT104 test strains. On top between control DT104 and post epidemic human strain S21. Below between control DT104 strain pre epidemic human strain S27 and post epidemic human strain S16. Prophage 5 from DT104 is highlighted in yellow and the *Salmonella* Genomic Island 1 is highlighted in green**



The lower part of Figure 4-10 shows the total absence of the SGI-1 region from two strains, S16 and S27, both of which were sensitive to the panel of antibiotics that the sequenced strain was resistant to.

Aside from the presence, absence or variation within SGI-1, the insertion in S21 and the atypical strain S15, the optical mapping technique revealed the largely clonal nature of strains from the phage-type DT104.

Discussion

The optical mapping technique demonstrated the ability to identify variation between strains. The comparison between the sequenced DT104 and the sequenced LT2 demonstrates that the variation lies largely with the phage and prophage regions and the SGI1. The other region shown, between prophage 1 (ST104) and Fels-1 is the allantoin / glyoxylate region also identified by the CGH study. The variation observed in the SGI1 regions of the optical maps for strains S04 and S21 confirms that from the CGH and the antimicrobial MIC.

Novel insertions were observed with the optical maps, a 40kb insert was identified in strain S21, between the genes *ppk* and *guaB.* The gene *ppk* codes for a polyphosphate kinase and the disruption of its transcription may have implications for the fitness of the strain. This information and the results from the microarray suggest the ubiquitous nature of phages within *Salmonella* Typhimurium, and infer a role for them in the evolution of DT104. The possible cost of phage acquisition on the fitness of the affected strain can also be investigated through the Biolog phenotyping study.

5 Phenotyping microarray results

Introduction

The previous sections have focused largely on the genetic differences between the test strains that span the DT104 epidemic and other reference strains, notably *S.* Typhimurium LT2, a sequenced strain. These studies have shown a number of differences that are likely to impact on the phenotype of the organism: the presence of allantoin and glyoxylate utilisation genes, the absence of a ribose operon for one strain and the *pdu/cob* operon for another. There is also the possibility that a novel insertion, which may be a phage, in S21 may have affected the transcription of phosphate kinase. However the methods have their limitations and for CGH, the key limitation is the statistical algorithms to establish cut-offs for gene presence and absence. Whilst being confident that the analyses are sound, it is important to verify any gene differences observed by CGH by alternative means.

The obvious test for those gene changes from the LT2 backbone that impact on phenotype is to verify the phenotype. To do this a study using the Biolog phenotyping microarray was used. This works by measuring the respiratory conversion of a tetrazolium dye by the bacteria in the presence of sole carbon, nitrogen, phosphorus or sulphur compounds in a minimal medium. The technique therefore allows the screening of the strains against a panel of compounds that represent those that may be available for central metabolism, and as such can give a measure of the fitness of the strain and a reflection of its genetic content.

5.1 Novel phenotyping approach using Biolog Phenotypic MicroArray™

The overall results for the phenotypic array carried out for the test strains are presented in Appendix 2 of the electronic submission. For this the cut-off for positive respiration was an OD of 150 at the 24 hour point. Statistical grouping of the metabolic results via Pearson correlation did not group them by date of isolation, host species or resistance profile.

As a baseline comparison the sequenced DT104 strain uses 82 of the 95 carbon compounds on PM1, 8 of the 95 carbon compounds on PM2a, 31 of the 95 nitrogen compounds on PM3, 34 of the 59 phosphorus compounds and 10 of the 35 sulphur compounds on PM4a. A comparison between test strain compound utilisation and the sequenced DT104 at the 24 hour point is shown in Table 5-1. Overall the strains are between 71.61% and 93.75% similar in their metabolic profile to the sequenced DT104. With the removal of data from PM2a (as will be discussed below), the similarity does increase slightly from 73.96% - 94.44%. This can be compared with the range of genetic similarity, 99.57% - 94.27%, as shown by the CGH microarray

The data shows some broad patterns. Human strains use an average of 48.3 extra compounds compared to 13.4 in the animal strains. The variations within these numbers is notable, S15 using 109 compounds, S16 using 85 and S07 using 80. At the other end of the scale strain S21 uses only 10 extra compounds and strain S22 uses 11 extra.

For compounds not used, when compared to the sequenced DT104, animal derived strains show less utilisation than human strains. Human strains have a median

value of 6 compounds not used and average of 17.5, compared to 13 and 23.8 respectively for the animal strains. Some individual strains do show higher non-usage for example S21 not utilizing 68 compounds, S22 not utilizing 67, D03 not utilizing 53 and D06 not utilizing 61.

Table 5-1 **Comparison of compound usage between test strains and sequenced DT104 strain from the Biolog phenotypic microarray plates ant the 24hr point. Data presented as the percentage of similarity. The second percentage column shows the same data but with the results from plate PM2a excluded.**

Strain	Compounds in common with DT104	Additional compounds	Compounds not used	% similarity to DT104	% similarity, - PM2a	
S01	313	62	9	81.51	90.63	Human derived strains
S02	308	70	6	80.21	91.32	
S03	320	58	6	83.33	94.44	
S04	325	52	7	84.64	86.11	
S05	319	58	7	83.07	93.40	
S06	333	46	5	86.72	88.19	
S07	302	80	2	78.65	86.81	
S13	316	62	6	82.29	93.06	
S15	275	109	0	71.61	86.46	
S16	294	85	5	76.56	92.36	
S19	352	28	4	91.67	90.97	
S21	306	10	68	79.69	73.96	
S22	306	11	67	79.69	73.96	
S24	349	20	15	90.89	89.93	
S27	341	39	4	88.80	87.85	
S28	348	31	5	90.63	89.58	
S29	353	21	10	91.93	90.28	
S30	343	37	4	89.32	87.85	
A01	338	9	37	88.02	85.07	Animal derived strains
A02	344	5	35	89.58	86.46	
D01	354	17	13	92.19	90.28	
D02	352	24	8	91.67	89.58	
D03	328	3	53	85.42	81.25	
D04	358	5	21	93.23	91.67	
D05	346	11	27	90.10	88.54	
D06	321	2	61	83.59	78.82	
D07	343	28	13	89.32	87.15	
D08	352	18	14	91.67	89.24	
D09	355	19	10	92.45	90.63	
D10	349	28	7	90.89	88.19	
D11	360	14	10	93.75	92.36	
D12	356	4	24	92.71	91.32	

5.2 Carbon source plate PM1

On PM1 the sequenced DT104 was able to use 82 of the 95 carbon compounds. Human strains could on average use more compounds than strains of an animal origin, at 83 compared to 73. Animal strains A01 (58 compounds) and D06 (66 compounds) use the fewest carbon sources, while strains S15 (93 compounds) and S16 (92 compounds) use the most. Compounds that show different utilisation patterns are listed and specific examples are highlighted below.

5.2.1 m-Tartaric acid and Glyoxylic acid

Table 5-2 shows the positive utilisation of two compounds, m-Tartaric acid and Glyoxylic acid. Figure 5-1 shows the metabolism of Glyoxylic acid and Figure 5-2 that of m-Tartaric-acid. These graphs demonstrate the ability of post epidemic animal strain A02 and post epidemic human strains S15 and S16 to metabolise glyoxylate and post epidemic animal strain strains A02 and D06 and post epidemic human strain S15 to metabolise m-Tartaric acid.

Table 5-2 Pattern of m-Tartaric acid and Glyoxylic acid utilisation from the Biolog phentotypic microarray plates.

Compound	S01	S02	S03	S04	S05	S06	S07	S13	S15	S16	S19	S21	S22	S24	S27	S28	S29	S30	A01	A02	D01	D02	D03	D04	D05	D06	D07	D08	D09	D10	D11	D12
m-Tartaric Acid	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-	-
Glyoxylic Acid	-	-	-	-	-	-	-	-	+	+/-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-

Figure 5-1 **Reduction of tetrazolium violet dye by the 32 test strains incubated at 37°C over a 26 hr period in the presence of the sole carbon source Glyoxylic acid.**

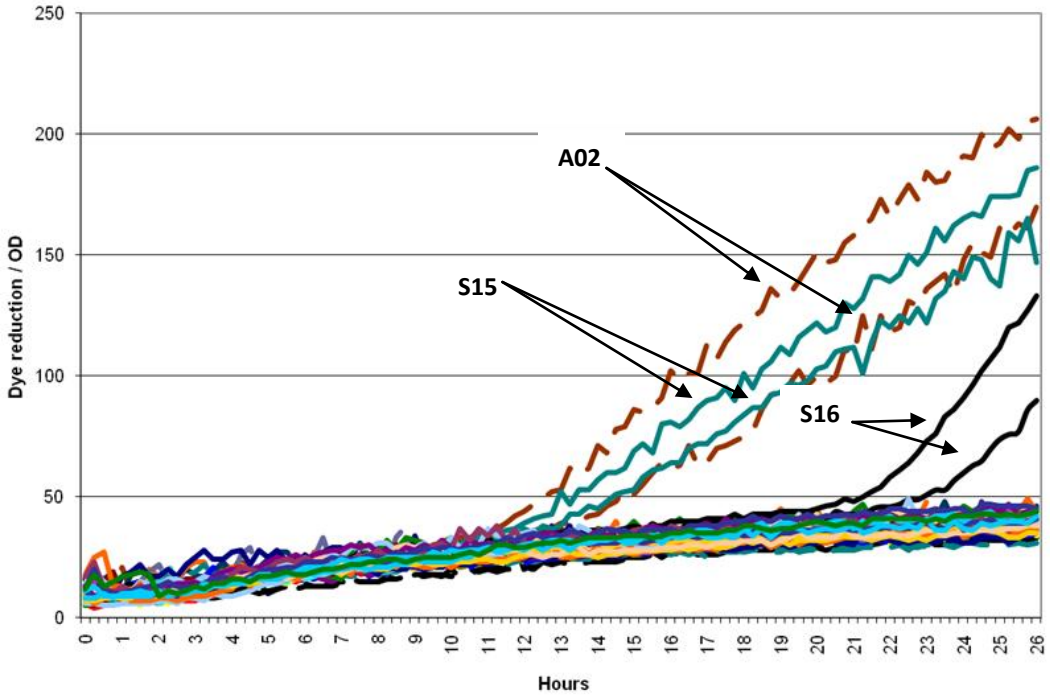
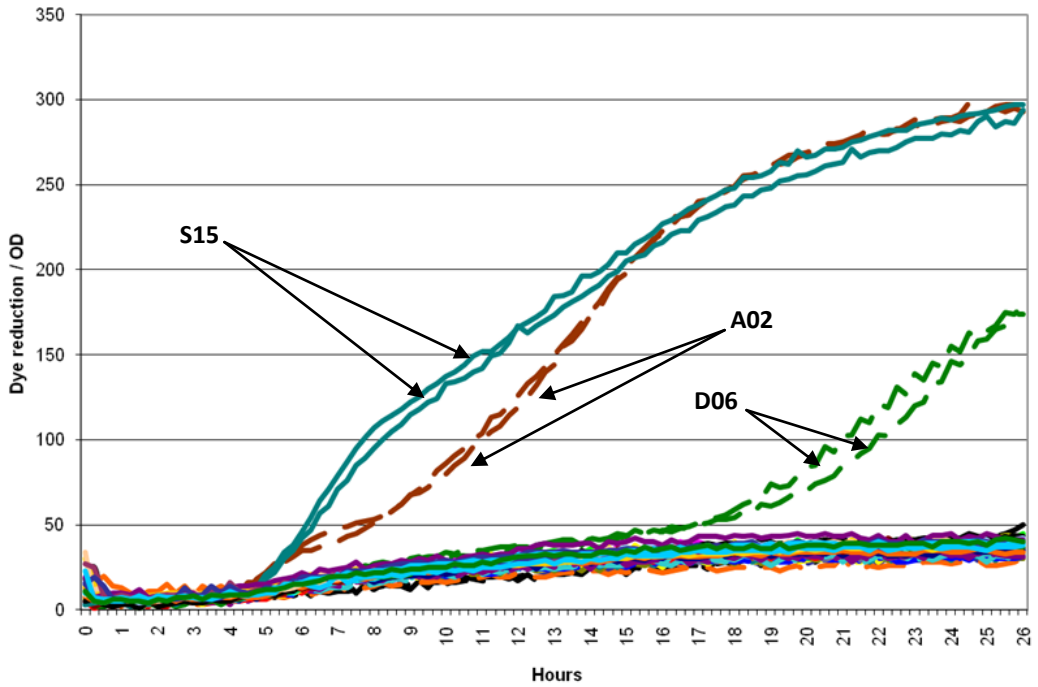


Figure 5-2 **Reduction of tetrazolium violet dye by the 32 test strains incubated at 37°C over a 26 hr period in the presence of the sole carbon source m-Tartaric acid**



The carbon compounds m-Tartaric acid and glyoxylate identified from above were further investigated via a growth rate assay to confirm the phenotype and directly connect increased respiration to growth. Figure 5-3 shows the results for the strains grown in minimal medium with 0.4% Glyoxylic acid, pH controlled. The control strains LT2 and *S.* Gallinarum and the test strain S15 all showed the ability to grow in the minimal medium with 0.4% glyoxylate w/v. No growth was seen with test strains S01, S04, S07, S21, S24, S27 and S30. Also no growth was achieved with the same growth media substituting glyoxylate with m-Tartaric acid with any of the test strains.

5.2.2 Tricarballylic acid

The utilisation pattern for the sole carbon source Tricarballylic acid is shown in Figure 5-4. Strain S16 was shown to have an extended lag phase and less respiration than the other strains.

5.2.1 N-Acetyl-D-Glucosamine

Figure 5-5 shows the pattern of growth with N-Acetyl-D-Glucosamine as a sole carbon source. The compound clearly differentiates human derived stains from animal derived ones. The upper group contains the human strains S01-S07, S13, S15-S16, S19, S21-S22, S24 and S27-S30 while the lower group contains the animal strains A01-A02 and D01-D12. A similar pattern, although not as clearly defined, exists for the compound Acetoacetic acid.

Figure 5-3 **Growth of strains in minimal medium with 0.4% glyoxylate w/v. Increased growth, shown by increase in optical density at 600nm against time. LT2, *S.* Gallinarum and S15 show log. growth at around 48 hours**

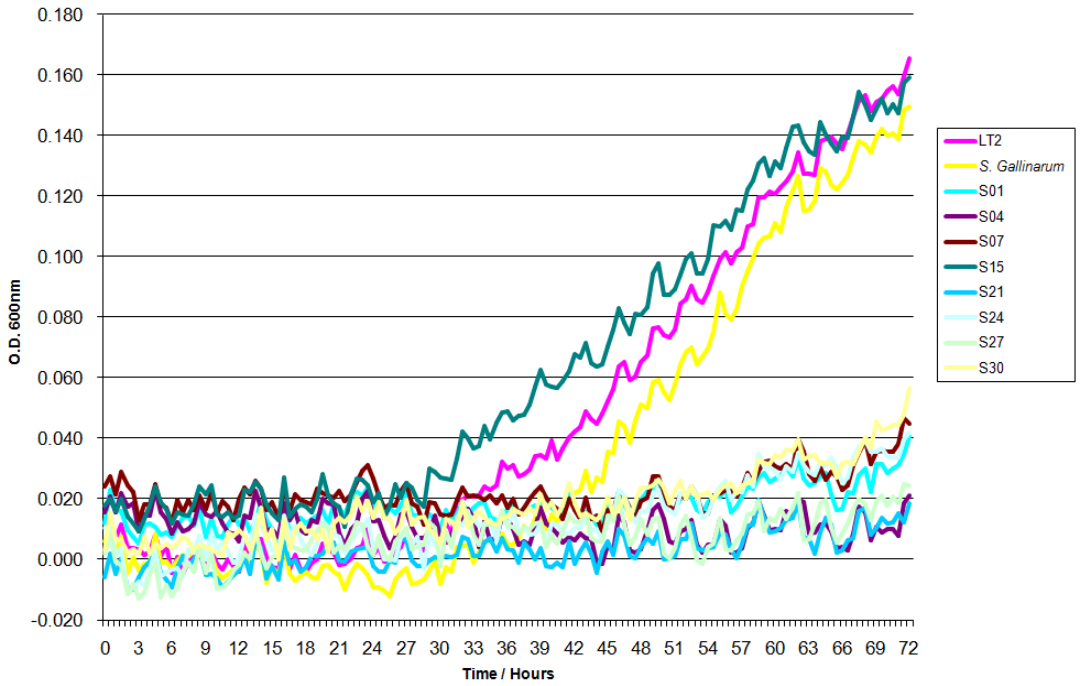


Figure 5-4 **Reduction of tetrazolium violet dye by the 32 test strains incubated at 37°C over a 26 hr period in the presence of the sole carbon source Tricarballylic acid**

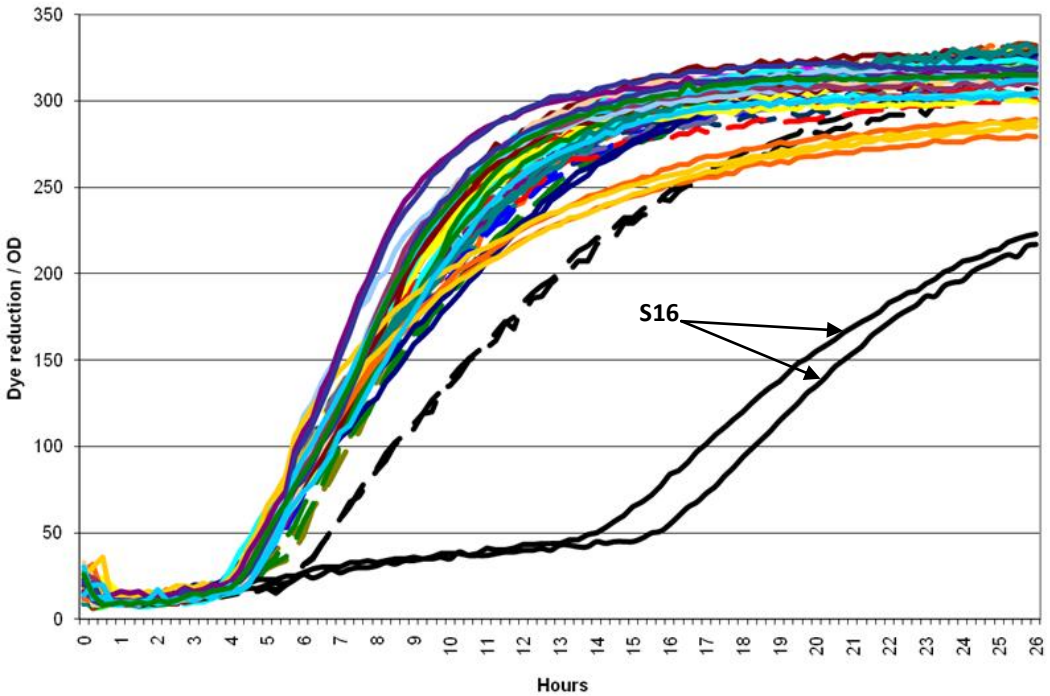


Figure 5-5 **Reduction of tetrazolium violet dye by the 32 test strains incubated at 37°C over a 26 hr period in the presence of the sole carbon source N-Acetyl-D-Glucosamine**

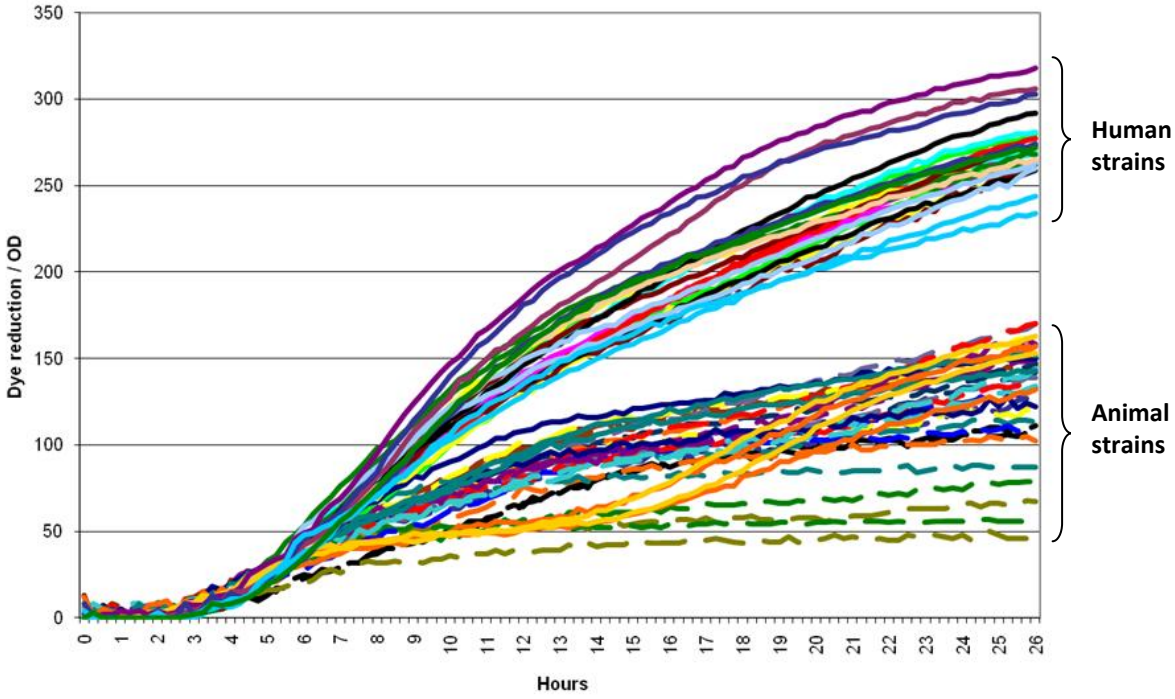
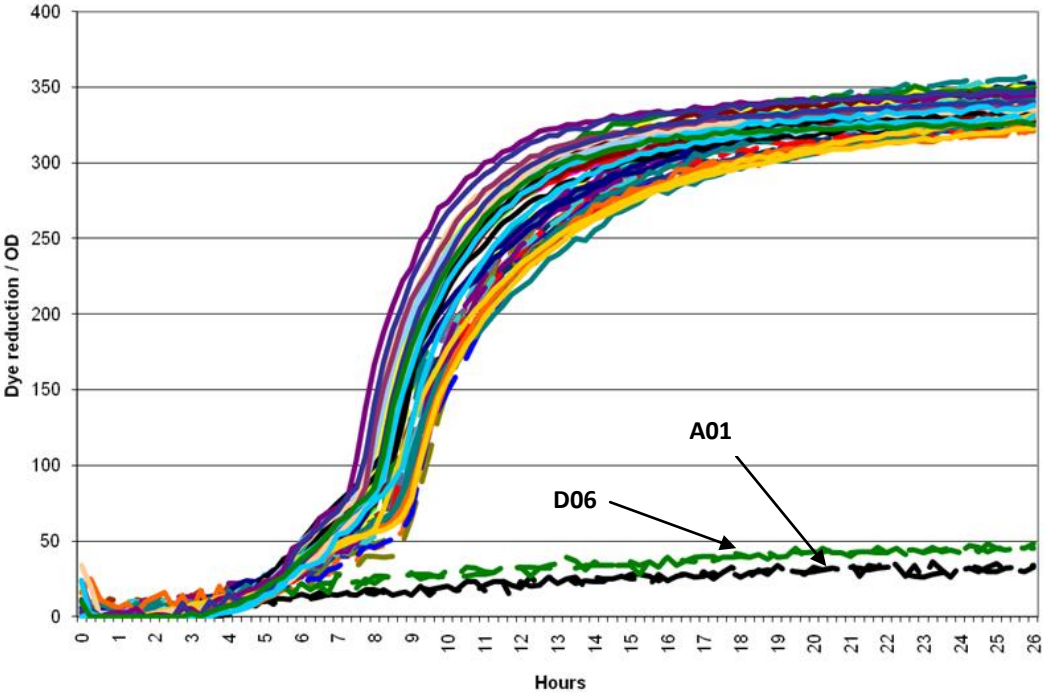


Figure 5-6 **Reduction of tetrazolium violet dye by the 32 test strains incubated at 37°C over a 26 hr period in the presence of the sole carbon source D-Saccharic acid**



5.2.2 D-Saccharic acid

Figure 5-6 shows that two animal strains A01 and D06 showed little or no respiration when D-Saccharic acid was the sole carbon source.

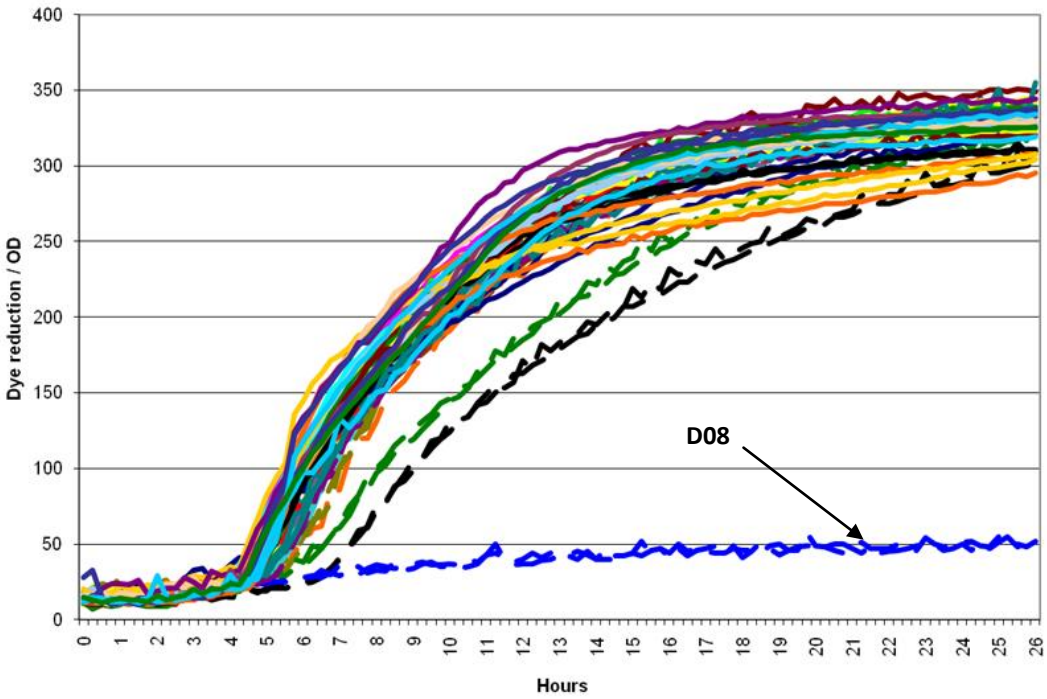
5.2.3 D-Cellobiose

Some human strains demonstrated comparatively better ability to use D-Cellobiose as a sole carbon source. This group, which includes strains S27, S28, S22, S21, S19, S16 and S30, showed increased respiration when compared to the other human and animal strains. One caveat is that the replicates for these strains showed a degree of divergence, although they did show a collective positive response. It is possible that this divergence may be due to a slightly different concentration of the compound within the original plates.

5.2.4 Tyramine

Figure 5-7 shows that one animal derived strain, D08, showed little or no respiration when grown in minimal media containing the compound Tyramine as a sole carbon source.

Figure 5-7 **Reduction of tetrazolium violet dye by the 32 test strains incubated at 37°C over a 26 hr period in the presence of the sole carbon source Tyramine**



5.3 Carbon source plate PM2a

On PM2a the sequenced DT104 strain shows the ability to use 8 of the 95 carbon compounds (AbuOun et al., 2009). The first ten strains tested with this plate, (S01-S07, S13, S15 and S16) showed positive results from between 25 to 78 compounds, whilst the remaining 22 strains showed positive reactions with 5 to 16 compounds. Re-testing with two strains from the first group, S04 (12 compounds) and S15 (10 compounds), suggested that the largely positive reaction from that group may be an experimental anomaly. Thus the results from PM2a should be treated with a degree of caution in relation to the strains in the first group. For that reason, fewer examples have been drawn from this data, and the ones that have been chosen are, at least in part, confirmed by the retested strains.

5.3.1 D-Lactic acid methyl ester

Figure 5-8 shows that animal strains A01 and D04 - D12 had no ability to use D-Lactic acid methyl ester as a sole carbon source. The retested strains S04 and S15 also fit this profile, both showing positive responses.

5.3.2 L-Sorbose

Strain S16 showed a greater ability to utilise L-Sorbose than the other strains (Figure 5-9). Other strains that responded did so in only one of their duplicates, and should be considered suspect. The work of Woodward and Charles (Woodward and Charles, 1982) showed that L-Sorbose was used variably amongst the *Enterobacteriaceae* and that some strains were fully capable, others mutable and other incapable. It would appear the same is true within this group of test strains.

Figure 5-8 **Reduction of tetrazolium violet dye by the 32 test strains incubated at 37°C over a 26 hr period in the presence of the sole carbon source D-Lactic acid methyl ester**

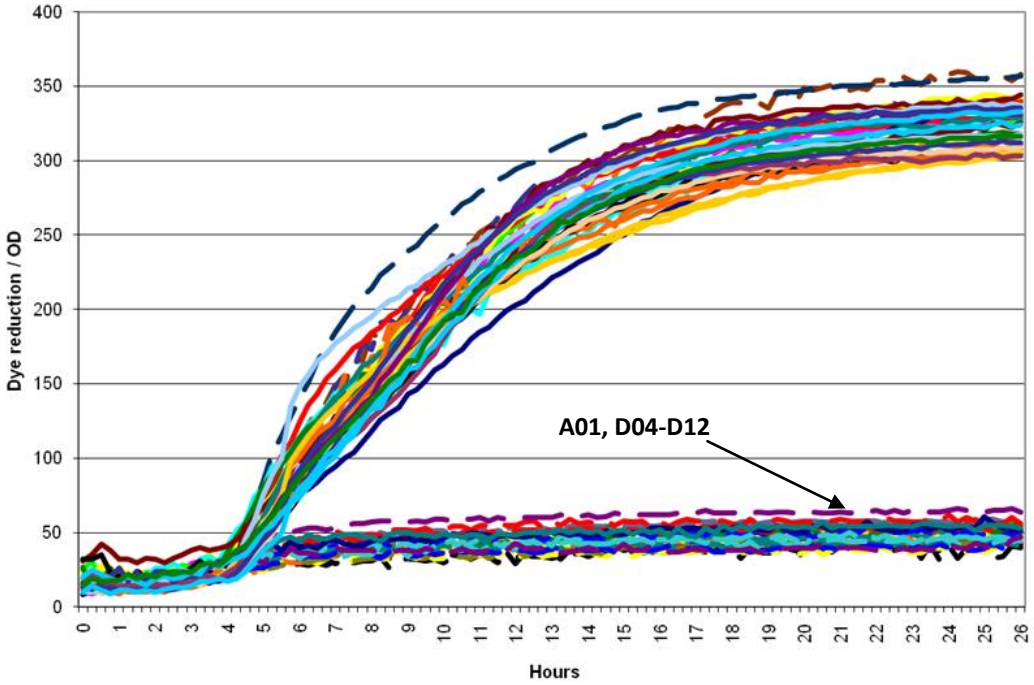
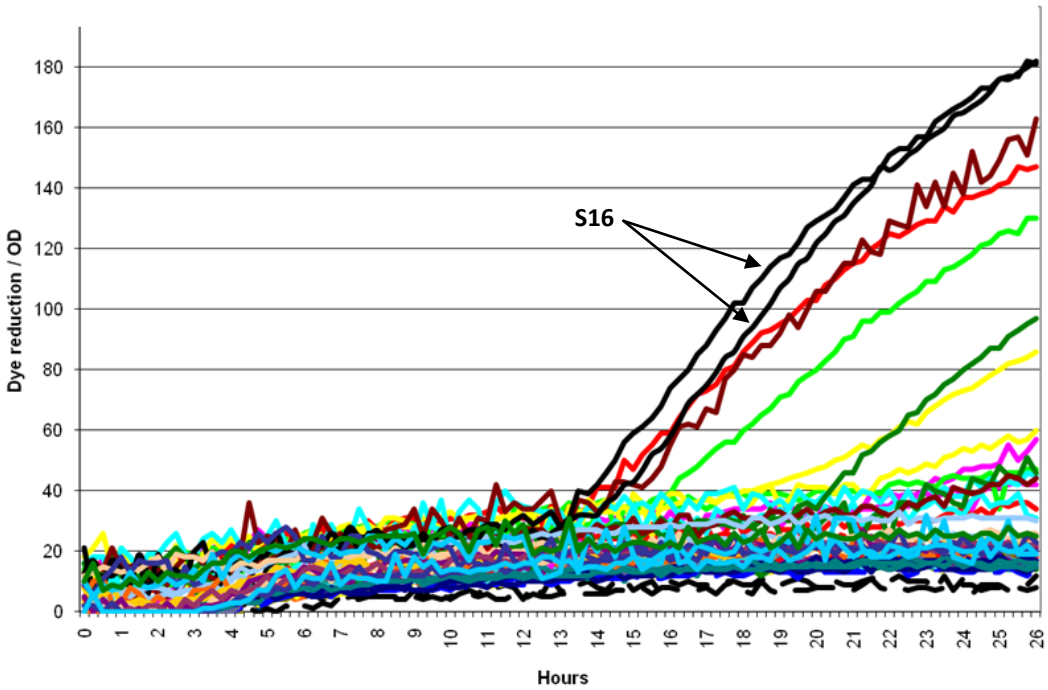


Figure 5-9 **Reduction of tetrazolium violet dye by the 32 test strains incubated at 37°C over a 26 hr period in the presence of the sole carbon source L-Sorbose**



5.4 Nitrogen source plate PM3b

Plate 3b contains 95 different nitrogen compounds and uses succinate as a carbon source. The sequenced DT104 strain is able to utilise 31 compounds as a sole nitrogen compound (AbuOun et al., 2009). Two human strains, S21 and S22, showed a marked reduction in the number of usable nitrogen compounds compared to the others. This suggests a fundamental problem with the uptake or metabolism of sole nitrogen sources. These two strains were limited to metabolism of uric acid and four dipeptide pairs; ala-asp, ala-gln, gly-met and met-ala. The other test strains were positive for between 23 and 45 compounds, with human strains having a slightly higher average than animal strains at 39 compared to 31.

5.4.1 Ammonia

Three human strains (S04, S07 and S27) showed the ability to utilise ammonia as a sole nitrogen source (Figure 5-10). The sequenced DT104 strain showed no growth when ammonia is the sole nitrogen source, so the growth of these three strains is notable. Some of the other strains began to reduce the dye towards the end of the time period, perhaps suggesting that eventually strains begin to adapt to the sub-optimal conditions.

Figure 5-10 **Reduction of tetrazolium violet dye by the 32 test strains incubated at 37°C over a 26 hr period in the presence of the sole nitrogen source ammonia**

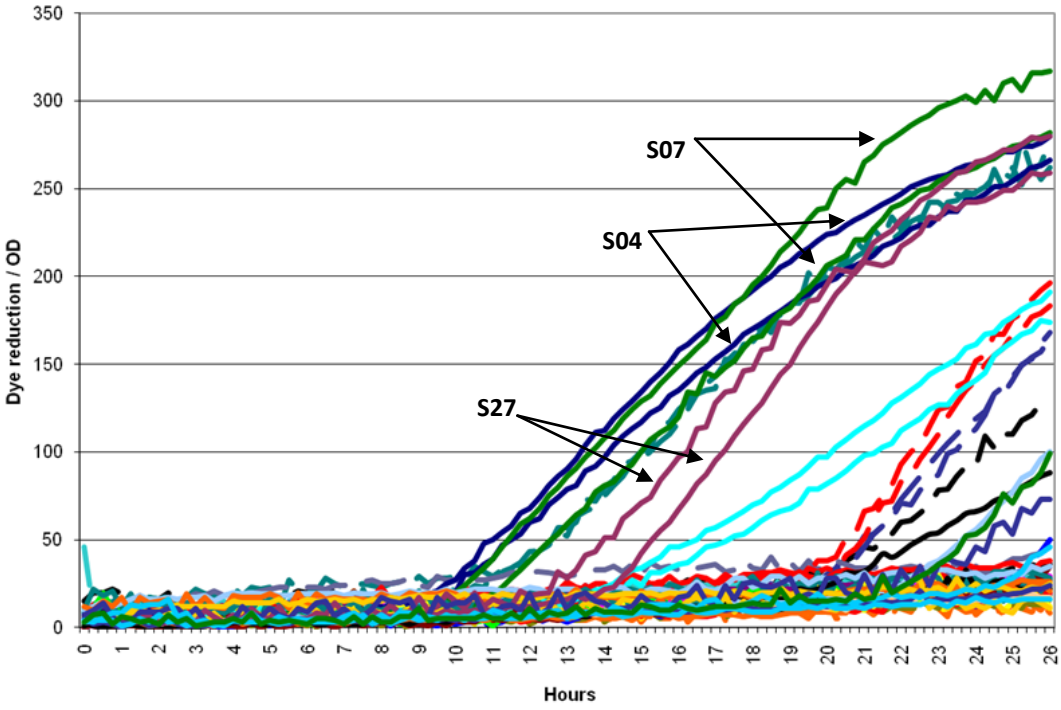
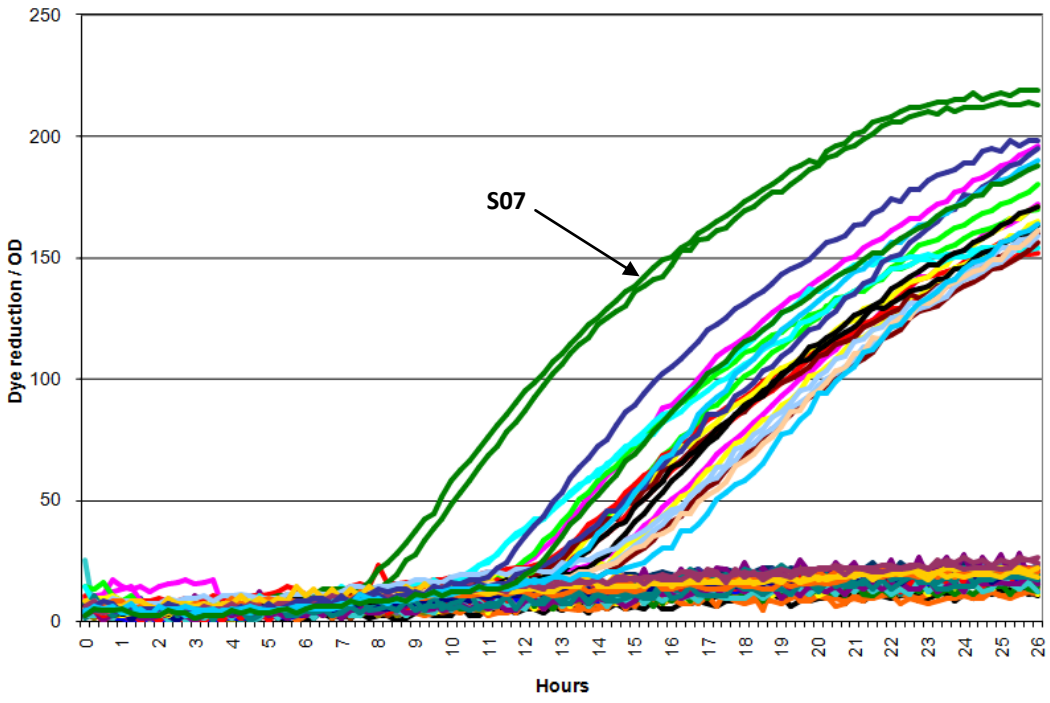


Figure 5-11 **Reduction of tetrazolium violet dye by the 32 test strains incubated at 37°C over a 26 hr period in the presence of the sole nitrogen source N-Phthaloyl-L-Glutamic acid**



5.4.2 N-Phthaloyl-L-Glutamic acid

With the nitrogen compound N-Phthaloyl-L-Glutamic Acid strain S07 was the first to show a significant respiratory response, followed by the other human strains except strains S21, S22 and S27 (Figure 5-11). These latter strains and all those from animal origin did not demonstrate respiration when this compound was the sole nitrogen source.

5.4.3 L-Ornithine and γ-Amino-N-Butyric acid

Strain S07 has shown the ability to utilise both ornithine and γ-Amino-N-Butyric acid (GABA) (Figure 5-12 and Figure 5-13). The use of ornithine as a sole nitrogen source during growth is dependent on the presence of the enzyme ornithine decarboxylase (ODC) (Shaibe et al., 1985b).

5.4.4 D-Mannosamine

The majority of human strains except S04, S15, S21 and S22, were able to use D-Mannosamine as a sole nitrogen source. The remaining human strains and the animal strains showed a much extended lag phase, with this compound and low levels of respiration in comparison with the other group (Figure 5-14).

Figure 5-12 **Reduction of tetrazolium violet dye by the 32 test strains incubated at 37°C over a 26 hr period in the presence of the sole nitrogen source L-Ornithine**

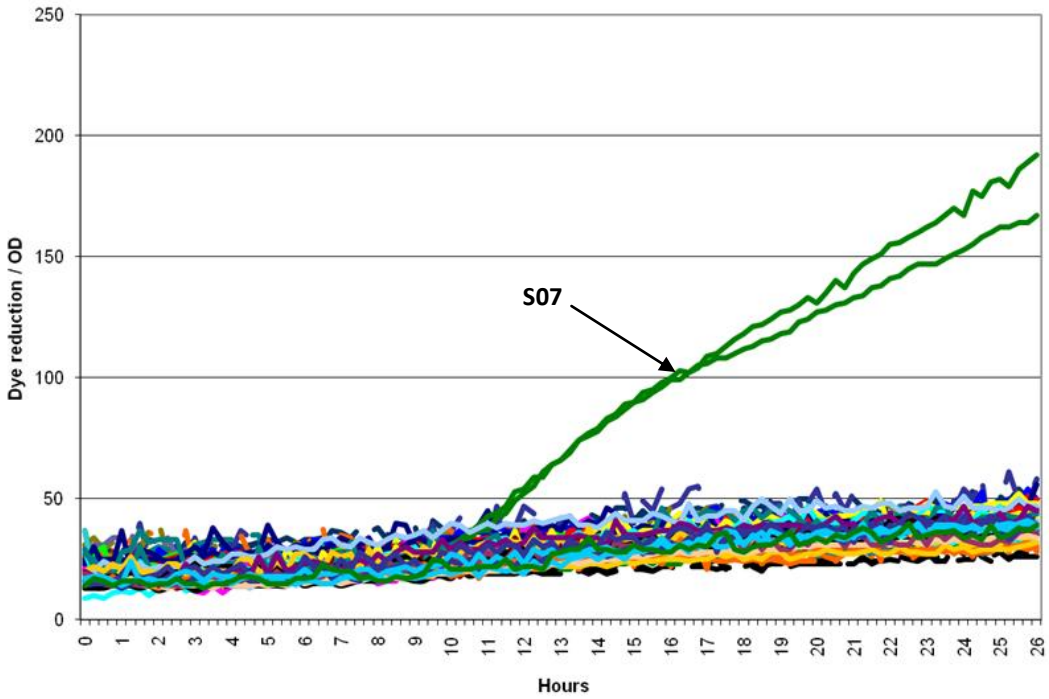


Figure 5-13 **Reduction of tetrazolium violet dye by the 32 test strains incubated at 37°C over a 26 hr period in the presence of the sole nitrogen source g-Amino-N-Butyric Acid**

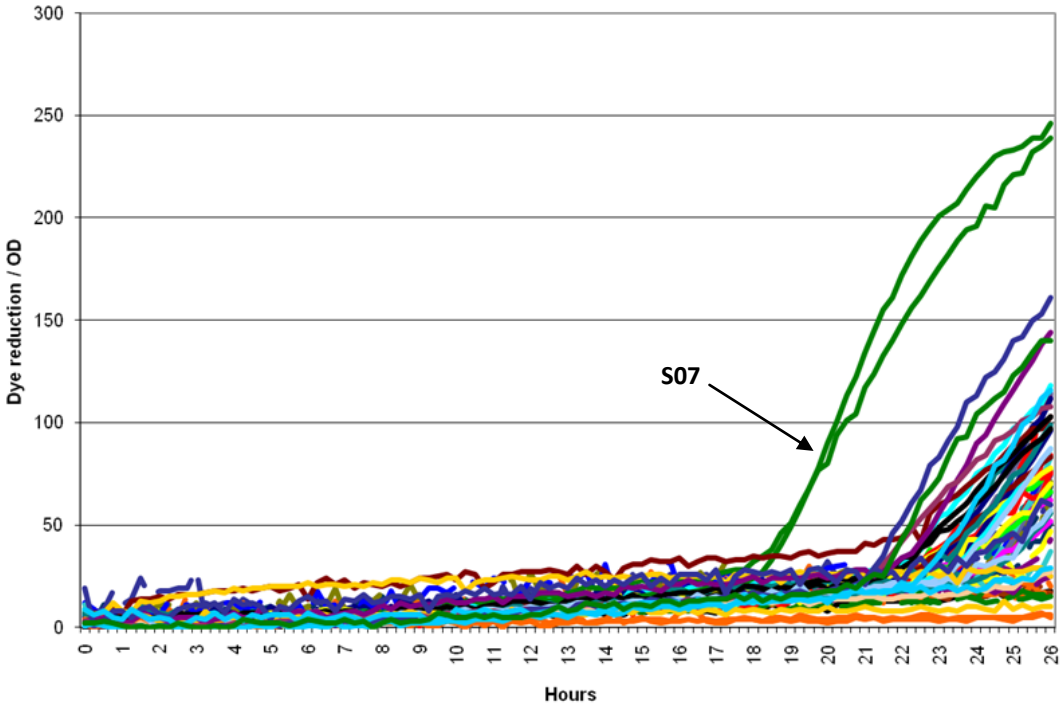


Figure 5-14 **Reduction of tetrazolium violet dye by the 32 test strains incubated at 37°C over a 26 hr period in the presence of the sole phosphate source D-Mannosamine**

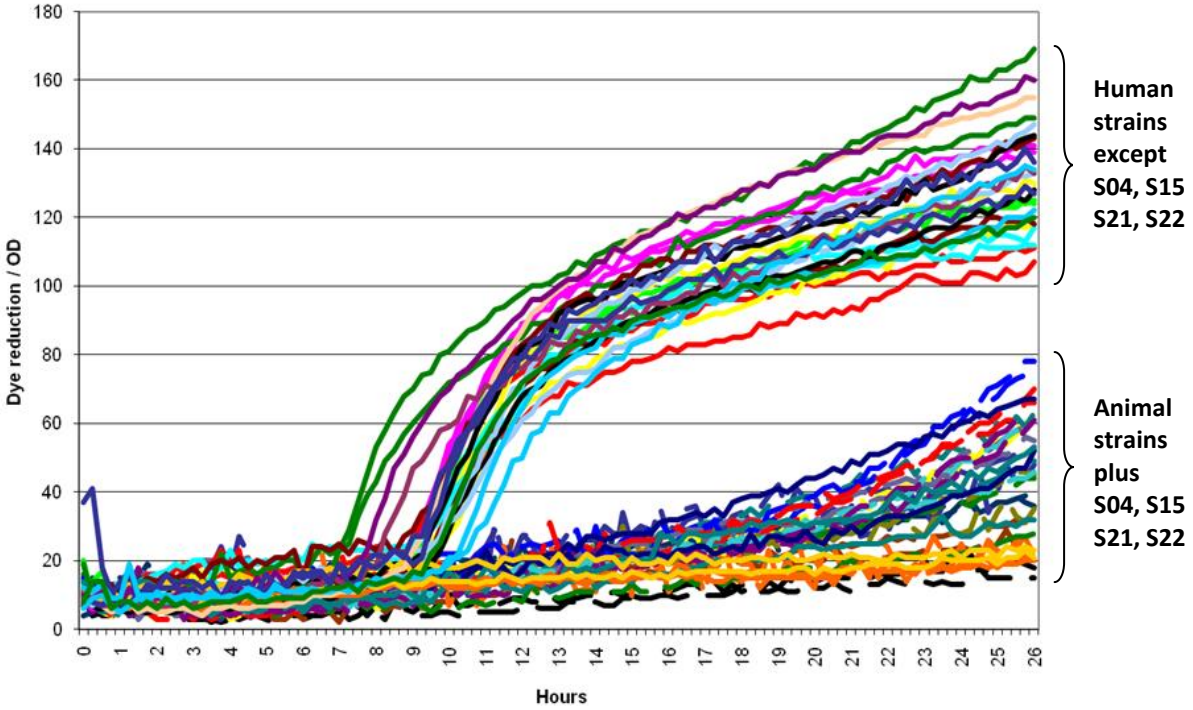
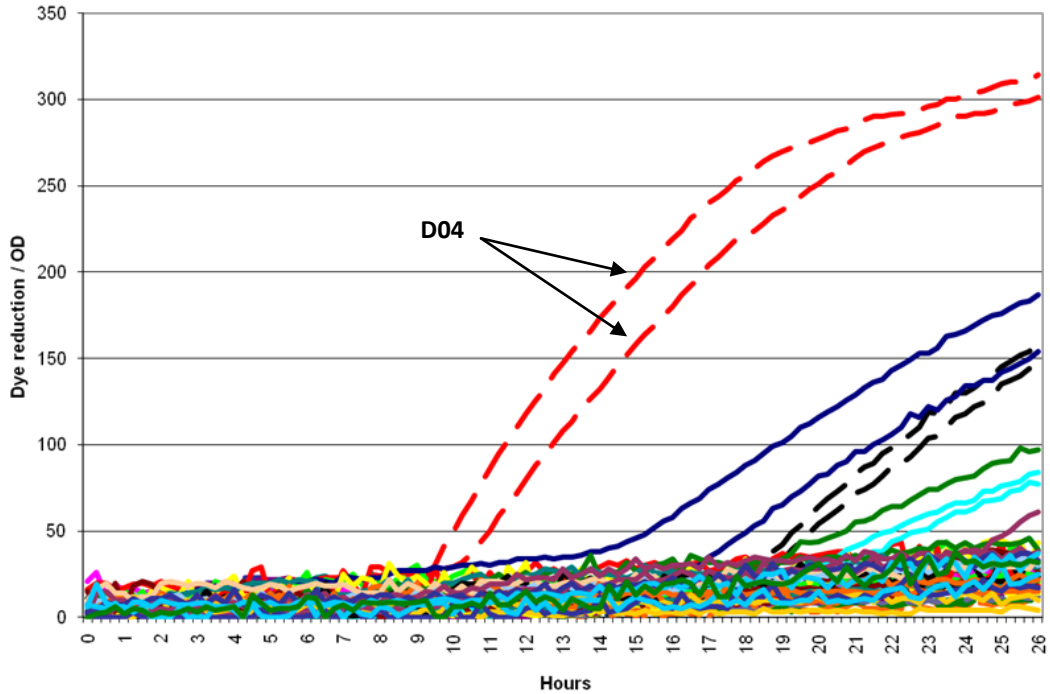


Figure 5-15 **Reduction of tetrazolium violet dye by the 32 test strains incubated at 37°C over a 26 hr period in the presence of the sole nitrogen source 6-Phospho-Gluconic acid**



5.5 Phosphorus source plate PM4a

Plate PM4a contains 59 single phosphorus sources and uses succinate as a carbon source. The sequenced DT104 strain uses 34 of these compounds (AbuOun et al., 2009). As with the nitrogen plate the most noticeable difference was the lack of utilisation by strains S21 and S22. They were only able to respire on two compounds, thiophosphate and dithiophosphate. Together with the results from the nitrogen plate this again suggests a fundamental problem with either uptake or downstream metabolism. The other strains range between 11 and 47 in the number of compounds they can use as a sole phosphorus source, with strain S04 using the most.

5.5.1 6-Phospho-Gluconic acid

Figure 5-15 shows the ability of animal strain D04 to use the compound 6-phosphogluconic acid as a sole phosphate source. 6-phosphogluconic acid has been shown not to be a sole carbon or nitrogen source for *S.* Typhimurium LT2 (Gutnick et al., 1969). The oxidative decarboxylation of 6-phosphogluconate to ribulose 5-phosphate by phosphogluconate dehydrogenase is part of the oxidative phase of the pentose phosphate pathway and results in the generation of NADPH from NADP+. This pathway allows the oxidation of glucose for the generation of NADPH, which is required for the synthesis of fatty acids, cholesterol and nucleotide biosynthesis.

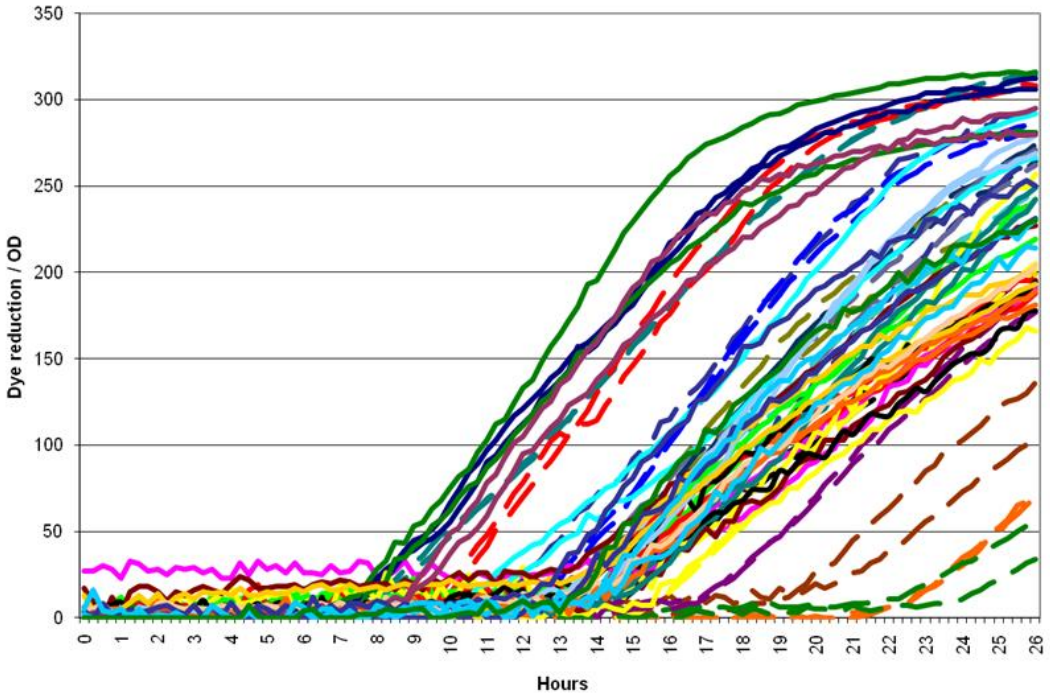
In *S.* Typhimurium LT2 the level of phosphogluconate dehydrogenase did not vary during variable growth conditions (Winkler et al., 1978), whereas in *E. coli* the level of 6-phosphogluconate dehydrogenase has been shown to be directionally proportionate to cellular growth rate during growth in minimal media (Jones et al., 1990).

5.6 Sulphur source plate PM4a

Plate 4a also contains 36 sole sulphur sources using succinate as the carbon source. The sequenced DT104 was able to utilise 10 of these compounds. Three animal strains A01, D03 and D06 showed the inability to use any of these compounds as sole sulphur sources. The rest of the strains used between 9 and 19 of these compounds.

A group of four strains showed a tendency for a shorter lag phase when utilising these compounds (S04, S07, S27 and D10). Figure 5-16 shows an example of this with the compound dithiophosphate acting as the sulphur source.

Figure 5-16 **Reduction of tetrazolium violet dye by the 32 test strains incubated at 37°C over a 26 hr period in the presence of the sole phosphate source 6-Phospho-Gluconic acid**



Discussion

The phenotypic microarray was carried out to help confirm the genetic variations that were revealed by the CGH and the optical mapping studies. The microarray results have confirmed some of the previous studies' results, disagreed with others and have also posed further questions and possibilities with regards to factors that may affect epidemicity.

The presence of the genes involved in allantoin and glyoxylate utilisation were reported in strain S15, A01, A02 and D06 by the CGH microarray. The phenotyping results support this observation for strains S15, A02 and D06 which showed the ability to utilise glyoxylate as a sole carbon source as well as m-tartaric acid which is also part of the glyoxylate metabolic pathway. This ability was confirmed and definitively linked to growth with S15 via a growth study. The ability to use these compounds, which is hypothesised to have become lost before DT104 obtained antibiotic resistance, is likely to have an effect on the fitness of the strains in the wild.

The absence of the ribose operon from one strain, S27, was also investigated with the phenotypic array. In this case the genotype could not be associated with a corresponding phenotype as all strains show the ability to use the compound (graph not shown). Mutations in this region have been shown to disrupt the transport system encoded but do not necessarily lead to no growth with D-Ribose as a sole carbon source, suggesting an additional transport mechanism for the compound. An intermediate in the L-Fucose degradation pathway, L-Fuculose-1-phosphate, has been

shown to stimulate the utilisation of D-Ribose via the ribose operon (Autieri et al., 2007)

One marked example of where the genetic information from the optical mapping may be linked to an extreme phenotype is with strain S21. As stated the 40kb insert identified between the genes *ppk* and *guaB*, may affect the production of a polyphosphate kinase. *Ppk* mutants have been shown to have reduced survival and sensitivity to weak organic acids. Strains S21 and S22, both from the same clinical outbreak, (the latter of which was not submitted for optical mapping) show dysfunction in the use of the large majority of sole nitrogen and sulphur sources. It is possible that the novel insertion has impacted to reduce their fitness in limiting conditions or has deactivated some vital process in nitrogen and phosphorus metabolism.

The other variation in the metabolism of sole compounds cannot be directly linked to specific gene differences observed from the CGH study. For example the absence of the *pdu/cob* genes from strain D10 was not mirrored by an alteration of this strains respiration response on 1,2 propanediol. Another example was the utilisation of the sole nitrogen source ornithine by strain S07. As has been stated the use of ornithine as a sole nitrogen source during growth is dependent on the presence of the enzyme ornithine decarboxylase (ODC). In *S.* Typhimurium two ODC genes, are present *speC* (STY3270) and *speF* (STY0739) - but only as pseudogenes, and *S.* Typhimurium has been shown to be unable to use ornithine as a sole nitrogen source (Gutnick et al., 1969).

An interesting observation from the phenotypic results was the divergence in utilisation between animal and human strains in the use of N-Acetyl-D-Glucosamine

as a carbon source as well as the majority of human strains in the use of D-Mannosamine as a nitrogen source. The main difference noted from the CGH between the human and animal groups was variations in the presence of genes related to phages.

Overall some aspects of the phenotyping can be linked to genetic differences from the CGH, however the majority cannot. In those cases where genes are present but the phenotype is absent, the implication must be that genes are inactive, possibly through mutation or their regulatory machinery. The inactivation of some genes, perhaps by cold stress, may be responsible for those instances when individual strains show different utilisation profiles from the majority of the test strains.

5.7 Persistence within macrophages

Introduction

The previous sections have focused on attempts to define the genetic and biochemical differences and similarities between the strains covering the time span of the DT104 epidemic. The data suggest a degree of clonality within DT104 strains and that the major variability is related to SGI1 and certain prophages as well as with some strains the region related to allantoin and glyoxylate utilisation. The Biolog data however suggests there were more subtle variations and phenotypic variability than might be anticipated on the assumption that the CGH data will correlate directly with the phenotype. Whilst that correlation is true the majority of genes, no doubt there were some discrepancies suggesting variability within genes, such as point mutations or alternative regulatory mechanisms, resulting in phenotypic differences.

A key part of the thinking around the emergence of an epidemic is the possibility that the bacterial strains themselves, through adaptation, become more competitive in the 'wild' and cause the rise of an epidemic. Similarly the decline of an epidemic may be the result of loss of fitness of the bacterial strains. Thus, given the differences that have been described in the panel of strains by the preceding descriptive genotypic and phenotypic tests that have been performed thus far in the study, it seemed sensible to test a facet of the biology of the strains that relates more to the relationship with the host. The induction of the glyoxylate cycle has been shown to be induced inside macrophages and so a study into the persistence of a panel of strains inside macrophages was carried out.

5.7.1 U937 human macrophage results

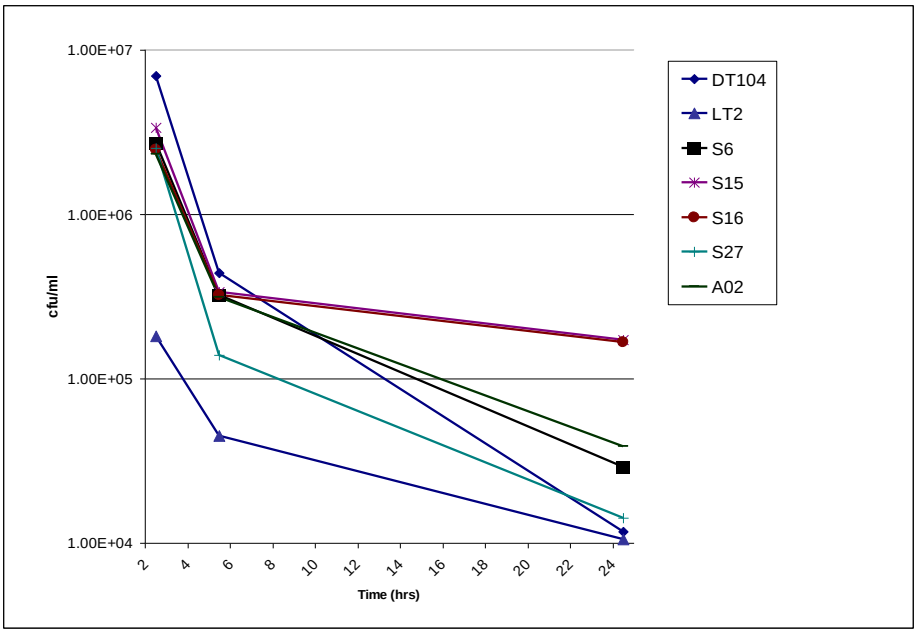
The survival and even proliferation of bacteria in the intracellular spaces of the host, be they in the epithelial cells of the gut or within macrophages, will have an important impact on the outcome of a *Salmonella* infection. U937 human macrophage and RAW 264.7 murine macrophage cells were chosen for the assay, and the protocols derived from (McNally, 2006) and (McNally, 2007) were followed. Representative strains were chosen with varying resistance profiles. Persistence was measured as cfu/ml after incubation with gentamicin after either 5 or 24 hours.

Figure 5-15 shows the data obtained from the seven strains at the three time points (2 hour, 5 hour and 24 hours). At the control point the DT104 strains show comparable numbers, with the sequenced DT104 strain being around half a log. higher. Each of the DT104 strains shows around a 1 log drop in numbers from the control point until the 5 hour persistence point, with S27 having a slightly larger drop.

The main difference between the DT104 strains is shown at the 24 hour persistence time point. Here the sequenced DT104 strain showed a large decrease of almost 2 logs, while the strains S06, S27 and A02 showed about a 1 log drop-off, and strains S15 and S16 showed a very small decrease indicating a higher level of persistence.

A variance analysis was carried out to assess the significance of the data. These included a comparison between all strains at all time points and between just the DT104 strains at all time points. The data for the DT104 strains is shown in Table 5-3. The null hypothesis that there is no significant difference in the variance between the strains at each of the time points is rejected with a confidence over 99.5%.

Figure 5-15 **Persistence in U937 macrophages. Data obtained from seven of the strains at the three time points; 2 hours (control, represent adhesion and invasion) and 5 and 24 hours (representing persistence). Each point shows the average of the replicates.**



Graph plotted on a logarithmic scale. Error bar represent the standard error (standard deviation divided by the square route of the number of samples.

Table 5-3 **ANNOVA values for DT104 strains at the three time points. All F values exceed the F-critical value for P values <0.05. Thus variance between strains is statistically significant.**

	F	*P-value*	*F crit*
Control (2hrs)	4.330	0.001	2.272
5hrs	2.490	0.034	2.272
24hrs	3.091	0.011	2.272

6 Discussion

Non-typhoidal *Salmonella* infections are a major public health concern. There are over 2500 serovars of *Salmonella* but a relative few of them cause the majority of human infections (Anjum et al., 2005). When an epidemic strain emerges it can spread across a wide geographic region before it then subsides. This pattern has been seen with successive epidemics for example with *S.* Enteritidis PT4 (Velge et al., 2005) and multidrug resistant *S.* Typhimurium DT104. MDR DT104 which was first isolated in the UK in the 1980's, was later found to have spread to Canada, France, Germany and the United States. First found in cattle, DT104 was later isolated from poultry, pigs and sheep as well as from human infections. As has been mentioned earlier in this thesis, it is recognised that an epidemic occurs through a combination of factors that can be simplistically considered as the interaction between the pathogen, the host and the environment. The genetic mechanisms underlying this epidemic and their possible implications for the fitness of the strains were the target of this study.

The hypothesis for this study was that DT104 emerged due to genetic changes that contributed to enhanced fitness. To test this proposition isolates from before, during and after the epidemic were studied, to discern those genetic changes associated with the epidemic. This assumes that genetic differences play a key role in the overall process and that these may be distinct from genetic drift say through the accumulation of single-nucleotide polymorphisms (SNPs). The evidence for this hypothesis is the fact that before and after the epidemic, fully sensitive strains of DT104 were commonly isolated whereas during the epidemic, MDR isolates were most often recovered. DT104 had indeed acquired a significant number of genes that were not just antibiotic resistance genes but were part of an associated element with

44 ORF's genes defined as the *Salmonella* Genomic Island 1. The gaining of antibiotic resistance by a clone of DT104 that became epidemic may have triggered the epidemic, and the variants and loss of the acquired genes could have reverted the strains to a 'less epidemic' state and thus led to the decline of the epidemic. The alternative could be that a clone of DT104 acquired the new genes and has since evolved separately. In this case one could argue that for DT104 the fully sensitive clone is 'less epidemic' than the MDR derivative. After the epidemic the less epidemic clone may still persist. For either of these scenarios it is assumed that the environmental and host contributions are broadly the same.

The strains for this study were chosen based on the criteria of date of isolation, host species and resistance profile. The *S.* Typhimurium serovar is commonly isolated from cattle but has also been found in other species. Therefore a comparison between strains isolated from animal and human sources enabled the consideration of genomic host specificity as well as epidemicity. In addition a range of antibiotic resistances have been reported in DT104, therefore variant profiles were chosen to allow consideration of the evolution of antibiotic resistance within the phage type, and its possible implications for the 'epidemicity' of DT104. Epidemicity is defined as 'the quality of being widely diffused and rapidly spreading throughout a population', (Saunders Comprehensive Veterinary Dictionary, 3 ed). Here the term is used in the context of increased isolations from a wider source of hosts. For DT104 it seems likely the primary source was bovine in nature (Threlfall, 1994), and other species first contracted it through food borne transmission or direct contact.

While considering the genetic underpinning of epidemicity within DT104 we should first consider if the CGH approach is justified. Preliminarily we should consider whether there are differences to consider at a genetic level beyond that of the known SGI1. The general clonality of the DT104 phage type has been reported on, however the presence of variant clones has also been observed (Liebana et al., 2002). Lan et al. (Lan et al., 2009) have looked into the population structure, origins and evolution of major *Salmonella enterica* clones using publically available MLST data. From the available *S.* Typhimurium MLST data they saw that the 46 isolates they studied were grouped into 12 strain types. Of these 9 belong to a clonal complex with ST19 as a founder, differing by between 1-5 base changes. A total of 93% of isolates reside within the ST19 complex. The remaining ST's are hypothesised as having independent origins, and these include ST36.

Within the *S.* Typhimurium serotype, DT104 is not genetically homogenous and was associated with 4 ST types (ST19, ST40, ST153 and ST209). Of these only ST40 lays outside the ST19 clonal complex. MLST data from the current study shows that 75% of the DT104 strains (26/30) reside within the ST19 complex, whilst only 4 strains; S15, A01, A02 and D02 were not designated as ST19.

These results may indicate that these outlier DT104s have either diverged from the main grouping of DT104, or that DT104 may have developed from one of these variant strains and that that strain may be considered an ancestral type. The mobility of phages could also suggest that two evolutionarily unrelated strains could have been infected with the same pattern of phages, which independently produced the DT104 type.

The typing techniques used also demonstrated their limitations, in the ability to detect further variation within the closely related DT104 strains. This justified the choice of comparative genomics to interrogate the underlying reasons for the development of epidemicity, by considering specific gene variation.

When studying genetic differences within DT104 to identify what characterises epidemic strains, it is then important to consider three factors; antibiotic resistance, fitness and virulence. These concepts will colour the interpretation of the data. The factors are interlinked and can be inter-dependent, and can therefore form a complex picture and have a profound impact on epidemicity

6.1 Antibiotic resistance

Antibiotic resistance is the first of the three strain features discussed here as being potentially associated with epidemicity. The emergence of *S.* Typhimurium DT104 as an epidemic strain in cattle in the late 1980's and early 1990's seems to have coincided with the strains acquiring antibiotic resistance, predominantly the penta resistance ACSSuT phenotype. The question must thus be posed, did the acquisition of antibiotic resistance by DT104 lead to it becoming an epidemic strain, or did a strain with the potential to cause an epidemic then gain antibiotic resistance, and has later variation or loss of the region led to the decline of the epidemic.

It has been shown that with Typhimurium DT104, resistance is chromosomally located on a 43kb genomic island (Boyd et al., 2001). This genomic island called SGI1 has been classified in the *Salmonella* Typhimurium DT104 and been shown to contain a 13kb region with genes encoding antibiotic resistance. The region is chromosomally located at the 3' end of the *thdF* gene, and is flanked by

18bp direct repeats. The region consists of a *floR* gene for chloramphenicol resistance and *tetR* / *tet(G)* for tetracycline resistance with two class integrons, one containing an *aadA2* cassette and the other a bla_{PSE-1} cassette (Briggs, 1999; Lawson, 2004; Mulvey et al., 2006). This represents the 'classical' genomic island from penta -resistant DT104 and confers the ACSSuT type.

The origin of this island is not known. In discussion with experts in the field (Rob Davies, personal communication), there was no evidence of major cattle husbandry changes at the time of the emergence of the epidemic as a contributing host or environmental factor. That said, the acquisition of the MDR phenotype may have given selective advantage in the face of antimicrobial treatments. It has been postulated that it emerged in response to the use of antibiotics during intensive calf rearing in the 1970's. The resistances conferred by the island target 4 of the 5 most commonly used veterinary antibiotics during this period (aminoglycosides, β-lactams, sulphonamides and tetracyclines (Velge et al., 2005)). Whether it was this that caused the creation and acquisition of the SGI1 or more subtle changes in antibiotic use around the epidemic period is unknown. Some of the genes; *aadA2*, bla_{PSE-1} and *sul1* are found amongst the *Enterobacteriaceae* whilst it has been suggested that the *floR* and *tet(G)* may come from the fish pathogens *Photobacterium damselae* and *Vibrio anguillarum*, where the genes have been found (Angulo et al., 2000). Another possible source for these later genes is *Pseudomonas* sp. that can contain. *tet(G)* (Schnabel and Jones, 1999) and a close homolog of *floR*, *cmlA*. These issues need further investigation but were considered significantly beyond the scope of this thesis.

Other variants of the genomic island have been reported with different resistance profiles, due to rearrangements or additions and / or subtractions to this

region (Levings, 2007; Threlfall, 2004). This change in resistance profiles shows the plasticity of this genetic system and the speed at which such changes can occur. The CGH microarray data has revealed variation within the SGI1 between strains. It has identified 4 strains (D02, S21, S22 and S24) as have the variant SGI1-C, although sequencing would be needed to confirm this. The strains that are sensitive, from both animal and human origin, lack all but three of the genes from this region, including those genes with phage/plasmid relatedness and those containing gene cassettes that confer the antimicrobial resistance.

The two genes shown to be present in all the test strains, regardless of antimicrobial resistance, were SDT3852 and SDT3862. Both have relatedness as determined by blast search to *qacEΔ1* and *sul1*, a gene pair associated with the 3' conserved region of a class 1 integron. Class 1 integrons have been associated with antimicrobial resistance in many Gram negative organisms and are characterized by the presence of a 5' conserved segment containing an integrase gene (*intI1*), a 3' conserved segment containing *qacEΔ1* and *sul1* genes, and a central attI recombination gene (Fluit and Schmitz, 1999). It is likely that positive presence of the two genes, even in strains that display no antimicrobial resistance, demonstrates the ubiquity of class 1 integrons. Indeed SDT3862 shows significant blast homology to part of the *S*. Typhimurium virulence plasmid pSLT.

Gene SDT3829 is also shown to be present in all the strains apart from A01, A02 and D06. This gene shows blast similarity to a *rep* gene and also to a mating pair stabilization protein gene. The *rep* gene from *E. coli* codes for a helicase, a DNA dependent ATPase needed by some phages for their replication (Gauss et al., 1994).

Therefore it is possible that the positive response to this gene from the majority of the sensitive strains may indicate its presence in another phage present in the genome.

Two of the strains used in this study presented expanded resistance profiles when compared to the 'typical' penta resistant DT104. Strain S05 demonstrated an additional resistance to trimethoprim. S06 showed additional resistances to the quinolone, naladixic acid and ciprofloxacin, a second generation fluroquinolone.

Quinolone based antibiotics work by targeting DNA gyrase or topoisomerase IV and therefore interfering with cell division. Resistance to these antibiotics is primarily conferred by mutations in the quinolone resistance-determining regions (QRDRs) of genes coding for these targets, *gyrA* and *gyrB* for DNA gyrase and *parC* and *parE* for topoisomerase IV. An additional role in resistance by efflux mechanisms associated with the AcrAB efflux pump has been suggested (Giraud et al., 2000). These resistance mechanisms are unrelated and distinct from the SGI1 region. This kind of mutational change would not be detected by the CGH microarray, and only gene sequencing techniques would reveal it. In addition, a secondary attachment for SGI1 has been discovered via R1 plasmid mediated transformation with *Salmonella* Typhimurium LT2. This is located in the intergenic region between the chromosomal *sodB* and *purR* genes (Doublet et al., 2008). This shifting of the island would not be detected by the CGH microarray, but could have an effect on the epidemicity of the strain if present.

Additional resistance genes, associated on the array with the *S.* Typhi plasmid HCM1, were identified in some of the resistant strains. These genes, ^{bla}*TEM*, *strA*, *strB* and *sul2*, have recently been associated with a clonal group of *S.* Typhimurium

and its monophasic variant *Salmonella enteric* sbsp. *enteric* serovar 4,[5],12:i:- that emerged during the 2000's. It has been isolated in Spain, Italy, Luxembourg and the US (Soyer et al., 2009). These strains display the resistance profile ASSuT and the resistance genes have been shown to occur in a separate and distinct chromosomally encoded region when compared to SGI1 (Lucarelli et al., 2010). This serotype is of increasing concern as it seems to be undergoing an epidemic phase. The presence of a subset of these HCM1 genes, but not the whole plasmid, suggests that the test strains from the current study may contain this newly identified region or another plasmid or plasmids with these genes.

The lack of information on the location of and spatial relationship between genes provided by the microarray results means that it is impossible to pin down the exact nature of the resistance profiles from this data alone. The ability of techniques such as sequencing and optical mapping to identify variation in SGI1 related regions and detect novel insertions demonstrate their ability to monitor the evolving nature of chromosomally based antibiotic resistance.

Overall the acquisition of the 'classic' resistance profile did coincide with the understood period of epidemicity. The later development of variant SGI1 regions also seems to have coincided with the decline of the epidemic, perhaps suggesting that the penta resistance profile acquired was the important factor in the prevalence of DT104. However some of these variations led, in fact, to resistances to a wider spectrum of antimicrobials. This may suggests that resistance in itself was not the only driving force of increased fitness and therefore the epidemic.

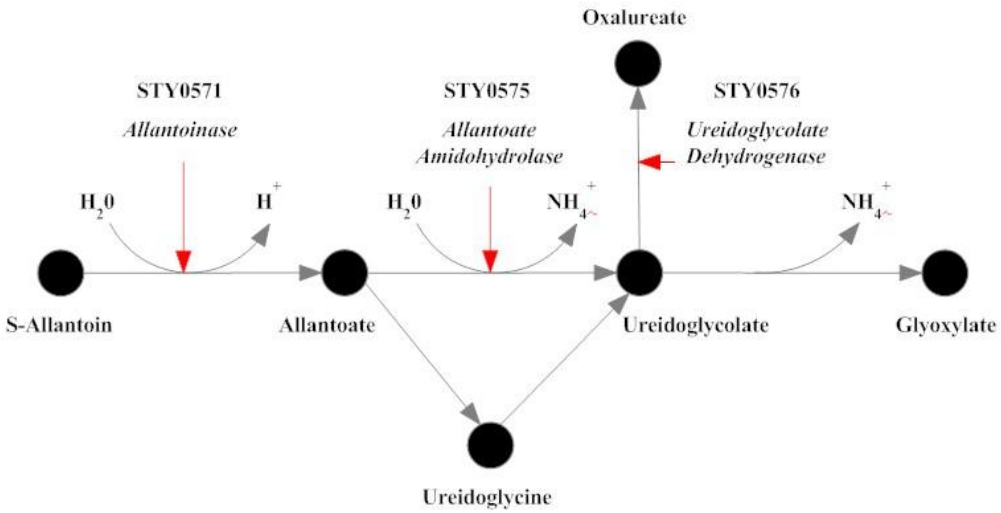
6.2 Fitness

Genomic differences that could affect fitness and therefore the spread of the strains - the second feature potentially related to epidemicity in this discussion - were identified by the microarray. Genes for allantoin and glyoxylate utilisation were found as being present in three sensitive animal strains, A01, A02, D06 and one of human origin, S15. Three of these strains – A02, D06 and S15 – showed the ability to utilise glyoxylate as a sole carbon source. The loss of this region of genes has been hypothesised to have occurred before antibiotic resistance acquisition (Matiasovicova et al., 2007), and as such, it is possible that a sensitive DT104 strain containing this region could represent an ancestral strain.

Allantoin is a product of purine degradation as it is transformed into urea and glyoxylate via ureidoglycolate. *E. coli* has been shown to use allantoin as a sole nitrogen source, but not a carbon source, under anaerobic conditions (Cusa et al., 1999). This pathway and the position of the genes within it are shown in Figure 6-1. Allantoin is broken down by successive enzymes, producing ammonia and ultimately glyoxylate. Glyoxylate can then be utilized in general metabolism. The process of allantoin degradation is part of a larger pathway that allows the breakdown of uric acid. Higher mammals, including humans, do not express uricase, the enzyme that breaks down uric acid to allantoin. It is therefore likely that allantoin will be unavailable in these hosts. In other host animals, including poultry, the enzyme is present and therefore allantoin may be available (Fujiwara and Noguchi, 1995). Allantoin is one of the materials that accumulates inside avian eggs during development and as such may provide one of the few nitrogen sources (Fisher and Eakin, 1957). The ability to use allantoin as a nitrogen source may be a significant

advantage to bacteria in these environments when other sources of nitrogen are scarce. The lack of the ability to utilise allantoin as a sole nitrogen source, by these strains, as determined in the Biolog study, may be an artefact due to the aerobic conditions of the assay.

Figure 6-1 Allantoin degradation pathway to Glyoxylate. Adapted from (Cusa et al., 1999).

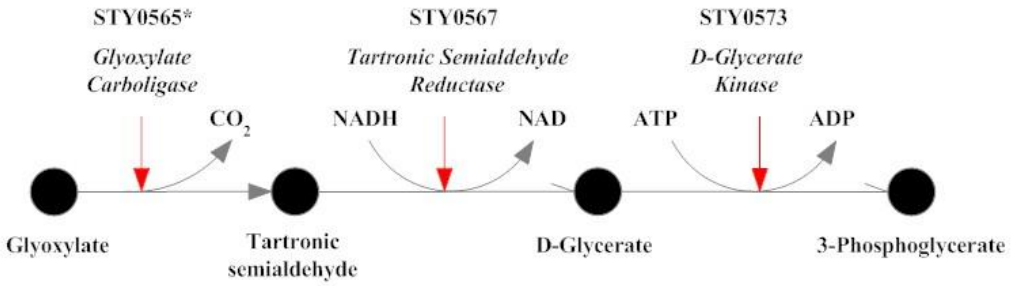


The glyoxylate cycle of central metabolism allows the bypassing of two decarboxylation steps of the TCA cycle, moving from isocitrate via glyoxylate to malate. It facilitates the use of 2-carbon compounds such as acetate and ethanolamine as sole carbon sources (Weissbach, 1953a, b). This process is termed the glyoxylate shunt and has been associated with growth in carbon limited environments. The genes present in the STY0566 – STY0577 that are involved in this process are identified in Figure 6-2.

The STY0567, *glxR* gene encodes for tartronic semialdehyde dehydrogenase, STY0573, *glxK* for glycerate kinase, and STY0566, *hyi* for hydroxypyruvate isomerase which interconverts tartronic semialdehyde and hydroxypyruvate

(Ashiuchi, 1999). The 3-phosphoglycerate formed can be fed into glycolysis metabolic pathway, and thus to central metabolism. It is suggested that the glycerate formed from glyoxylate is oxidised into the TCA cycle via acetyl-CoA in *E. coli* (Ornston and Ornston, 1969).

Figure 6-2 Glyoxylate degradation pathway to 3-Phosphoglycerate. Adapted from (Ornston and Ornston, 1969)



One of the suggested advantages of the presence of the ability to use the glyoxylate shunt of the TCA cycle is intracellular survival when nutrients are limited, for example inside macrophages (Lorenz, 2002). With a lack of more usual sources of carbon, such as glucose, the TCA cycle, with its two decarboxylation steps is unable to work. The glyoxylate pathway bypasses decarboxylation thus allowing 2-carbons to be used. The source of these 2-carbon compounds could be the breakdown of fatty acids from lipids within host cells (Chung, 1988; Maloy, 1980).

The presence of the genes involved in the allantoin / glyoxylate pathway would be thought to increase the fitness of the strain in some environments, specifically those that may be nutrient deficient. Whether the loss of this region may have had influence on either epidemicity or DT104 gaining antibiotic resistance is unknown. One speculative hypothesis could be that a reduction in the ability to

survive within intracellular spaces may have caused the bacteria to become exposed to more selective antimicrobial pressure in the extracellular environment.

The ability of strains to use other extra compounds compared to the sequenced DT104 may also confer a fitness advantage. For example the resistant strain S07 was observed to use ornithine as a sole nitrogen source. The *speC* gene for ornithine decarboxylase is found in four members of *Enterobacteriaceae; Citrobacter freundii*, *S.* Typhimurium, *Klebsiella pneumoniae* and *Enterobacter aerogenes* (Wright and Boyle, 1984). In *S.* Typhimurium two ODC genes, *speC* (STY3270) and *speF* (STY0739) were present but only as pseudogenes and *S.* Typhimurium has been shown to be unable to use ornithine as a sole nitrogen source (Gutnick et al., 1969). This is as opposed to the situation in *E. coli* where ornithine can be used (Shaibe et al., 1985a).

The pathway of ornithine utilisation leads to the formation of γ-amino-N-butyric acid (GABA). It has been shown that mutations in the *gab* regulatory locus lead to a decrease in the activity of GABA aminotransferase. Strain S07 also showed the ability to use GABA as a sole nitrogen source. These phenotypes may confer a fitness advantage to this strain. This result, from the phenotyping microarray, again demonstrates one of the drawbacks of the CGH microarray. A gene presumed 'pseudo' but actually actively transcribed will not be detected.

The ribose operon *rbsDABCKR* codes for a high affinity ABC transporter which allows the utilisation of D-ribose as a carbon source. It was identified on the *E. coli* K-12 chromosome at min 84 (Lopilato et al., 1984). The loss of this region in sensitive strain S27 is notable. It is possible that its excision is an artefact of lab

attenuation, perhaps because of its storage at low temperature. However, as has been noted, the region involved in allantoin and glyoxylate utilisation was lost from DT104 at some point before the epidemic and so perhaps this result demonstrates that even the regions outside the 'mobile elements' are more fluid then is generally thought.

As has been noted from the results from the optical mapping method a 40kb insert was identified in strain S21, between the genes *ppk* and *guaB.* If the *ppk*, polyphosphate kinase, gene has been disrupted there may be implications for the fitness of the strain. Mutant *ppk* derivative strains have been shown to have reduced survival and sensitivity to weak organic acids (Brown and Kornberg, 2004; Price-Carter et al., 2005). As was stated this strain and S22, likely from the same outbreak, show the same inability to use the majority of nitrogen and phosphorus compounds. Variation in the uptake of peptides by *S.* Typhimurium with mutations in the peptide uptake mechanisms; oligopeptide permease (OPP) dipeptide permease (DPP) and tripeptide permease (TPP) have been demonstrated and it is possible that a dysfunction could exist in one or all these mechanisms in the two strains (Goodell, 1987; Hiles, 1986). These strains are resistant human isolates from the post-epidemic period. The suggested decrease in their fitness had not prevented them from being isolated from human infections, but may be an indication of the process by which an epidemic strain begins to decline.

6.3 Virulence

Virulence is the third key strain feature discussed here as presenting a possible link between the genetic properties and epidemicity. Here, variation in a region associated with a 'pertussis like' toxin was observed in the DT104 strains through the CGH

microarray data. *Bordetella pertussis* produces an ADP-ribosyltransferase toxin that causes the ADP-ribosylation of the G proteins and interferes with intracellular communication. This class of toxin consists of two components: an 'A' unit coding for the toxin including an active site, and a 'B' unit required for receptor binding and translocation across cell membranes. ADP ribosyltransferases are associated with severe human diseases such as diphtheria, cholera and pertussis.

The genes that code for the toxin, listed as STY1361-STY1367 in the *Salmonella enterica* serovar Typhi CT18 genome, have been reported as inactive in DT104 because of at least one frame shift (Hermans, 2005a; Parkhill, 2001). However a recent study has shown that the genes in this region can be transcribed in DT104 and produce an ADP ribosyltransferase that can affect Chinese hamster ovary cells in a similar way to an *E. coli* shiga toxin (Uchida, 2009). A study using the selective capture of the cDNA transcribed sequences from macrophages containing *Salmonella* Typhi has reported that the genes STY1361-STY1367 are upregulated, which suggests that they play or may have played a role in bacterial virulence or survival within phagocytes (Faucher, 2005).

The pattern of presence of these genes in the DT104 strains is interesting. The human strains from the epidemic period, excepting S15, showed the presence of the genes, most of the animal strains and the early sensitive human strains showed the presence of some of the genes while five strains A01, A02, D02, D03 and S15 showed the complete absence of the region. The study into this region was extended, with confirmatory PCR and the sequencing of the 'toxin' genes. The absence of the region from the five strains was confirmed and variation in the region was seen amongst the other strains from the sequencing results, although it did not marry directly with the

CGH microarray data. If, through further study, they can be shown to be active in the strains from the epidemic period but not those from the pre-epidemic, it could have fundamental implications in understanding the mechanisms behind the epidemic. To draw such a conclusion much more extended investigation is required.

One of the overarching conclusions from the genetic data is that most differences can be connected to mobile genetic elements. These can be bacteriophage related, have plasmid origin, can represent genomic islands that can be shown to be mobilizable or are pathogenicity islands whose G+C content divergence from the general genome indicates their 'foreign' origin (for example SPI2 has G+C content of 43% compared to the rest of the *Salmonella* genome at 52% (Hensel, 2004)). The transfer and integration of these mobile elements allow *Salmonella* to take so called 'quantum leaps' in bacterial evolution (Groisman and Ochman, 1993).

The acquisition of pathogenicity islands is one of the main means by which *Salmonella* has obtained virulence determinants. *Salmonella* pathogenicity island 1, SPI-1, is present in *S. bongori* and all serotypes of *S. enterica* tested thus far (Ochman and Groisman, 1996). Sensitive strain S13 lacks three genes from this region; STY3006 *sipD*, STY3007 *sipC* and STY3014 *spaO*. *SipC* is a cell invasion protein associated with the needle-like apparatus that is expressed inside murine and porcine macrophages. *SpaO*, is one gene from the *inv-spa* invasion gene complex, coding for a component of the TTSS of SPI-1. SPI-1 is responsible for aspects of the invasion of non-phagocytic cells via the rearrangement of the host actin cytoskeleton and is also implicated in the inflammation of the intestinal epithelium (Hensel, 2004). Dysfunction in this area would likely have a negative impact of the pathogenicity and virulence of the bacterium.

The animal resistant strains D05, D10-D12 were shown to have a few additional genes from the identified SPI-7 region (Pickard et al., 2003). With D10 this includes the genes STY4617 and STY4622, two genes that bookend the genes *lysB*, *nucD* and *nucE* in SPI-7. The SPI-7 in *S.* Typhi contains the genes for the Vi antigen that is expressed as a capsular exopolysaccharide. The region also contained the *sopE* phage, which produces the effector protein SopE. The *S.* Gallinarum genome contains a degenerate SPI-6, that is present in sensitive strains A02 and S15, whilst the *S.* Typhi SPI-6 region is missing in the same strains. The deletion of SPI-6 from *S.* Typhimurium had no effect on systemic pathogenesis but does seem to lead to the reduced invasion of cultured cells (Folkesson et al., 2002).

Infection by bacteriophages may also provide a route for the acquisition of virulence determinants. As was observed the human strains, except S19 and S30, showed the presence of genes from the Fels-1 phage. A study by Hermans et al (Hermans, 2005b) on 18 *S.* Typhimurium DT104 strains did not detect the presence of the Fels-1 phage by genomic subtractive hybridisation. Some genes of Fels-1 have been shown to cross hybridise with open reading frames of prophage homologs in LT2 (Andrews-Polymenis et al., 2004). The presence of Fels-1 genes in some DT104 strains could have implications for virulence. Fels-1 contains *sodCIII*, a superoxide dismutase and *nanH*, a neuraminidase. Superoxide dismutase provides bacteria which some protection from the oxidative burst in the intracellular environment and is linked to bacterial survival. The role of neuramidase in virulence has not been elucidated, although a role for it in bacterial biofilm production during mucosal infection has been reported (Soong et al., 2006). The presence of these genes in the human DT104 strains could be indicative of the presence of this or a related phage.

Two animal resistant strains where shown to lack the Gifsy-1 phage. Gifsy-1 contains a *gogB* gene which produces a type III secreted substrate. This protein is secreted by the TTSS of the pathogenicity islands SPI-1 and SPI-2, and is translocated into the host cells in a SPI-2 mediated process, via the *ssrAB* system (Coombes et al., 2005). The *gogA* gene is also present in Gifsy-1, and shows homology to *pipA* from the SPI-5 of *S.* Dublin. Gifsy-1 also contains *ehly-1* which in enteropathogenic *E. coli* produces an enterohemolysin.

The presence of these mobile elements and the variation that has been shown within the DT104 strains demonstrates how strains obtain different combinations of virulence determinants. It also shows how a strain may suddenly develop characteristics that may both positively and negatively affect its epidemicity.

6.4 Conclusion

When considering epidemicity it is important to understand the factors that underlie the spread of a strain and its virulence. If a strain is too virulent it will so adversely affect its host that it limits its own spread, similarly if the acquisition of virulence determinants has a serious deleterious effect on a strain's fitness it again will not become epidemic. A highly fit and virulent strain with no antibiotic resistances would again be sorely hampered in becoming epidemic, being adversely affected by both man-made and naturally occurring antimicrobials. However the acquisition of antibiotic resistance may have same genetic cost on the fitness of a strain as the acquisition of virulence determinants.

So the barriers to epidemicity are significant, and yet wave after wave of epidemics are reported. Bacteria have developed two important strategies to enable them to overcome these obstacles. First, and not to be overlooked, is the growth rate of bacteria and the advantage this gives from an evolutionary standpoint. *E. coli* has a textbook doubling time of 20mins under ideal conditions, although this is no doubt extended in the wild. With the growth curves performed for this study, DT104 entered the exponential growth phase in LB-broth after 2 hours. This ability to proliferate rapidly means that, with the natural mutational frequency for bacteria between 10^{-7} and 10^{-11} per base pair per single round of replication (Madigan et al., 2006), mutations are quickly accumulated within the population. Many are of course reversed by cellular DNA repair mechanisms, or are deleterious and therefore evolutionary dead-ends. However the chance of a beneficial mutation becoming fixed is high enough to drive evolution in response to selective pressures. Indeed microbiologists take advantage of this process when they select for antibiotic resistance in laboratory strains. For example repeated growth of *Salmonella* strains on nalidixic acid gradient plates can select for stains that have developed resistance to the antibiotic, and it has been shown that this occurs because of spontaneous mutations in the *gyrB* gene (Yoshida et al., 1991).

The second process that bacteria use to facilitate their evolution and that can play a critical role in their epidemicity is horizontal gene transfer. As has been reported in this study the majority of genetic variation involves mobile or mobilizable genetic elements. This is the route by which *Salmonella* has gained both virulence determinants, in the form of pathogenicity islands, and antibiotic resistance, in the form of plasmids in *S.* Typhi and as genomic islands with DT104. In addition transduction or lysogenic conversion with bacteriophages containing virulence genes

can lead to these being integrated in the bacterial genome. The process of horizontal gene transfer allows bacteria to make evolutionary jumps, and may explain why a strain suddenly gains a set of genes that trigger its journey to becoming an epidemic strain.

The evolution of antibiotic resistance may in fact go hand in hand with that of virulence. One theory of the source of resistance genes is that they originate from organisms that produce antibiotics themselves, and act as a defence against their own products (Davies, 1997; Webb and Davies, 1994). Other resistances may have developed from genes with a different functional role within the genome. For example the AcrAB-TolC multidrug efflux system in DT104 mediates additional quinolone resistance (Baucheron et al., 2004). It has also been demonstrated that *E. coli* and *S.* Typhimurium extrude bile salts through the *acrAB* system, a prerequisite for the colonisation of the intestinal tract (Ma et al., 1995). *S.* Typhimurium mutants with a high susceptibility to bile salts were shown to have mutation in the *acrB* gene of the MDR efflux pump, and decreased ability to colonise the murine intestinal tract (Lacroix et al., 1996). This example shows that antibiotic resistance and virulence may develop together, and become co-selected.

In the case of the CGH microarray results for the DT014 strains it is difficult to pin down the precise reasons for the rise of the epidemic during the 1980's and 1990's. Part of this difficulty is due to the complexity of the interaction between antibiotic resistance, fitness and virulence and part is due to the drawbacks of the technique itself. It could be hypothesised that the presence of the SGI1 itself was the driving force behind DT104's rise, but this in itself would not necessarily lead to an increase in virulence and the clinical symptoms that would lead to its isolation. There

is a tantalising proposition that the presence of an ADP-ribosyltransferase in the epidemic human strains could have an effect, but this would have to be investigated much more intensely. The loss of the region involved in allantoin and glyoxylate metabolism may have changed the strain's preferred environmental niche to one that led it to become more virulent. Loss of metabolic genes has been linked to virulence, in *Shigella* the loss of the *cadA* for cadaverine sysnthesis increases its virulence, as that enzyme has been shown to inhibit the *Shigella* enterotoxin (Maurelli et al., 1998).

Overall the CGH microarray has identified variation between DT104 strains from the epidemic and non-epidemic periods. The assumption behind that analysis is that the CGH method generated data is 'accurate' and this should be treated with a degree of caution. On the most basic level the technique interpretation depends on the annotated sequences from a few control strains. Genes identified as present may in fact be pseudo, orthologs or paralogs and the reverse is also true - genes identified in the annotation as 'inactive' may be able to be transcribed in the test strain. Strain S07 and the *speC* genes for the use of ornithine may be an example of this. The variation with phage genes, which by their nature may have a high degree of degeneracy and mosaicism, is hard to draw definitive conclusions on.

The statistical processing of the CGH microarray data through the Genespring software using Lowess normalisation as well as the decisions taken on the use of the ln(Cy5/Cy3) ratio for the precise cut-off for present/absent genes, were made to minimise the possibilities of false positives. However one of the weaknesses of the CGH method is still its statistical robustness, and therefore an attempt should always be made to confirm observed differences. Since the location of 'spots' on the microarray is randomised, differences in sequential runs of genes, i.e. from the same

operon, as with the allantoin / glyoxylate utilisation region, can be considered sound, especially when supported by phenotypic data. Other gene differences such as with the SGI1 were confirmed with the MIC and the optical mapping. The 'pertussis toxin regions' presence and absence, because of its possible impact, was confirmed by PCR and further investigated by sequencing. The observation of other variation can be anecdotally confirmed by other studies, i.e. with variation in the *pdu/cob* region (Anjum et al., 2005). .

Given that the results from this study did show genetic differences with correlated and uncorrelated phenotypes for the test strains, it is now important to revisit the testable hypothesis and assess whether the changes were associated with epidemicity. This study has shown that non-epidemic early strains lack the SGI1 region and that the presence of this region is one of the defining characteristics of epidemic strains. The later variation of this region also corresponded with the time when the epidemic was in decline. The early non-epidemic strains along with strains with an animal origin show variation in a toxin associated region, whereas the epidemic strains have the complete region. Other DT104 strains with different ST types show the presence of genes involved in allantoin and glyoxylate metabolism, whose loss may have preceded the epidemic.

At first sight the major genetic differences uncovered by this study would be thought to have a detrimental effect on fitness and then by association on epidemicity. The insertion of a large genomic island into the genome, a dysfunction in allantoin and glyoxylate metabolism and even the possible production of a toxin all carry a cost for the bacteria and could be thought to reduce its fitness. However DT104 became an epidemic strain. In addition there were numerous other differences observed and it is

difficult to tease out whether these were ‘noise’ associated with ongoing evolution of the epidemic type or part of the genetic selection that was associated with ‘epidemicity’. This study had underlined the complexity of underlying bacterial epidemics but has also uncovered interesting avenues for further investigation which may lead to a better understanding of epidemicity.

6.5 Future work

There are many possibilities for future work given the plethora of data produced by the CGH microarray and the novel techniques used in this study. If this investigator were to continue, an examination of the importance and possible implications of the region coding for the ‘pertussis like’ toxin would have been the chosen direction. If the finding that it can produce an active toxin could be confirmed (Uchida, 2009), this could have serious repercussions for the understanding of the DT104 epidemic.

Such work would be carried out by the induction of the toxin from a panel of strains, with mitomycin C and hydrogen peroxide, followed by a RT-PCR study to detect the transcribed RNA. Further studies could then be done to identify if the toxin or its effect is detectable *in vitro* tissue culture and ultimately in an *in vivo* murine model.

The advances in both the cost and speed of whole genome sequencing mean that the usefulness of CGH microarrays is gradually waning. However it can still identify strains of interest, and therefore suggest candidates for more in-depth study. A number of strains from this study would be possible candidates for full sequencing. In addition the optical mapping technique used in conjunction with the CGH has

identified a novel insertion in one of the strains, that may have a deleterious effect on its phenotype, and this would be another target for sequencing. Another direction future work could take is to attempt to close the gap between the genetic and phenotypic data. This could be first done by repeating the Biolog assay under anaerobic conditions and then by the use of expression microarrays.

7 Bibliography

AbuOun, M., Suthers, P.F., Jones, G.I., Carter, B.R., Saunders, M.P., Maranas, C.D., Woodward, M.J., Anjum, M.F., 2009. Genome scale reconstruction of a Salmonella metabolic model: comparison of similarity and differences with a commensal Escherichia coli strain. *J. Biol. Chem*., 284: 29480-29488.

Ahmer, B.M., Tran, M., Heffron, F., 1999. The virulence plasmid of Salmonella Typhimurium is self-transmissible. *J. Bacteriol*., 181: 1364-1368.

Alcaine, S.D., Soyer, Y., Warnick, L.D., Su, W.L., Sukhnanand, S., Richards, J., Fortes, E.D., McDonough, P., Root, T.P., Dumas, N.B., Grohn, Y., Wiedmann, M., 2006. Multilocus sequence typing supports the hypothesis that cow- and human-associated Salmonella isolates represent distinct and overlapping populations. *Appl. Environ. Microbiol.,* 72: 7575-7585.

Alpuche-Aranda, C.M., Racoosin, E.L., Swanson, J.A., Miller, S.I., 1994. Salmonella stimulate macrophage macropinocytosis and persist within spacious phagosomes. *J. Exp. Med*., 179: 601-608.

Anderson, E.S., 1968. Drug resistance in Salmonella Typhimurium and its implications. *Brit. Med. J*., 3: 333-339.

Anderson, E.S., Ward, L.R., De Saxe, M.J., De Sa, J.D.H., 1977. Bacteriophage typing designations of S. Typhimurium. *J. Hygiene*., 78: 297-300.

Andrews-Polymenis, H.L., Rabsch, W., Porwollik, S., McClelland, M., Rosetti, C., Adams, L.G., Baumler, A.J., 2004. Host restriction of Salmonella enterica serotype Typhimurium pigeon isolates does not correlate with loss of discrete genes. *J. Bacteriol.,* 186: 2619-2628.

Andrews, J.M., 2001. BSAC standardized disc susceptibility testing method. *J. Antimicro. Chemo.,* 48: 43-57.

Andrews, J.M., 2006. BSAC standardized disc susceptibility testing method (version 5). *J. Antimicro. Chemo.,* 58: 511-529.

Angulo, F.J., Johnson, K.R., Tauxe, R.V., Cohen, M.L., 2000. Origins and consequences of antimicrobial-resistant nontyphoidal Salmonella: implications for the use of fluoroquinolones in food animals. *Microb. Drug. Resist.,* 6: 77-83.

Anjum, M.F., Marooney, C., Fookes, M., Baker, S., Dougan, G., Ivens, A., Woodward, M.J., 2005. Identification of core and variable components of the Salmonella enterica subspecies I genome by microarray. *Infect. Immun.,* 73: 7894-7905.

Ashiuchi, M., Misono, H., 1999. Biochemical evidence that Escherichia coli hyi (orf b0508, gip) gene encodes hydroxypyruvate isomerase. *Biochimica et Biophysica Acta.,* 1435: 153-159.

Autieri, S.M., Lins, J.J., Leatham, M.P., Laux, D.C., Conway, T., Cohen, P.S., 2007. L-fucose stimulates utilization of D-ribose by Escherichia coli MG1655 DeltafucAO and E. coli Nissle 1917 DeltafucAO mutants in the mouse intestine and in M9 minimal medium. *Infect. Immun*., 75: 5465-5475.

Baucheron, S., Tyler, S., Boyd, D., Mulvey, M.R., Chaslus-Dancla, E., Cloeckaert, A., 2004. AcrAB-TolC directs efflux-mediated multidrug resistance in Salmonella enterica serovar typhimurium DT104. *Antimicrob. Agents. Chemother.,* 48: 3729-3735.

Bäumler, A.J., Hargis, B.M., Tsolis, R.M., 2000. Tracing the Origins of Salmonella Outbreaks. *Science*., 287: 50-52.

Bertani, G., 1951. Studies on Lysogenesis. I. The Mode of Phage Liberation by Lysogenic Esherichia coli. *J. Bacteriol*., 62: 293-300.

Betancor, L., Yim, L., Fookes, M., Martinez, A., Thomson, N.R., Ivens, A., Peters, S., Bryant, C., Algorta, G., Kariuki, S., Schelotto, F., Maskell, D., Dougan, G., Chabalgoity, J.A., 2009. Genomic and phenotypic variation in epidemic-spanning Salmonella enterica serovar Enteritidis isolates. *BMC. Microbiol*., 9: 237.

Boyd, D., Peters, G.A., Cloeckaert, A., Boumedine, K.S., Chaslus-Dancla, E., Imberechts, H., Mulvey, M.R., 2001. Complete nucleotide sequence of a 43-kilobase genomic island associated with the multidrug resistance region of Salmonella enterica serovar Typhimurium DT104 and its identification in phage type DT120 and serovar Agona. *J. Bacteriol*., 183: 5725-5732.

Boyd, D.A., Shi, X., Hu, Q.H., Ng, L.K., Doublet, B., Cloeckaert, A., Mulvey, M. R., 2008. Salmonella genomic island 1 (SGI1), variant SGI1-I, and new variant SGI1-O in Proteus mirabilis clinical and food isolates from China. *Antimicrob. Agents. Chemother*., 52: 340-344.

Brenner, F.W., Villar, R.G., Angulo, F.J., Tauxe, R., Swaminathan, B., 2000. Salmonella Nomenclature. *J. Clin. Microbiol.,* 38: 2465-2467.

Briggs, C.E., Fratamico, P.M., 1999. Molecular characterization of an antibiotic resistance gene cluster of Salmonella Typhimurium DT104. *Antimicrob. Agents. Chemother.,* 43: 846-849.

Brown, M.R., Kornberg, A., 2004. Inorganic polyphosphate in the origin and survival of species. *Proc. Natl. Acad. Sci. USA.,* 101: 16085-16087.

Brüssow, H., Canchaya, C., Hardt, W-D., 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Bio. Rev.,* 68: 560-602.

Center for Disease Control. *Salmonella Surveillance: Annual Summary, 2006*. Atlanta, Georgia, US Department of Health and Human Services. Available from: http://www.cdc.gov/ncidod/dbmd/phlisdata/salmonella.htm [accessed 14 Oct 2012].

Chan, K., Baker, S., Kim, C.C., Detweiler, C.S., Dougan, G., Falkow, S., 2003. Genomic comparison of Salmonella enterica serovars and Salmonella bongori by use of an S. enterica serovar Typhimurium DNA microarray. *J. Bacteriol.*, 185: 553-563.

Chen, P., Andersson, D.I., Roth, J.R., 1994. The control region of the pdu/cob regulon in Salmonella Typhimurium. *J. Bacteriol*., 176: 5474-5482.

Chung, T., Klumpp, D.J., LaPorte, D.C., 1988. Glyoxylate bypass operon of Escherichia coli: cloning and determination of the functional map. *J. Bacteriol.*, 170: 386-392.

Cloeckaert, A., Schwarz, S., 2001. Molecular characterization, spread and evolution of multidrug resistance in Salmonella enterica Typhimurium DT104. *Vet. Res.*, 32: 301-310.

Cooke, F.J., Wain, J., Fookes, M., Ivens, A., Thomson, N., Brown, D.J., Threlfall, E.J., Gunn, G., Foster, G., Dougan, G., 2007. Prophage sequences defining hot spots of genome variation in Salmonella enterica serovar Typhimurium can be used to discriminate between field isolates. *J. Clin. Microbiol*., 45: 2590-2598.

Coombes, B.K., Wickham, M.E., Brown, N.F., Lemire, S., Bossi, L., Hsiao, W.W., Brinkman, F.S., Finlay, B.B., 2005. Genetic and molecular analysis of GogB, a phage-encoded type III-secreted substrate in Salmonella enterica serovar Typhimurium with autonomous expression from its associated phage. *J. Mol. Biol*., 348: 817-830.

Crosa, J.H., Brenner, D.J., Ewing, W.H., Falkow, S., 1973. Molecular relationships among the salmonelleae. *J. Bacteriol.*, 115: 307-315.

Crump, J.A., Luby, S. P., Mintz, E.D., 2004. The global burden of typhoid fever. *Bull. World. Health. Organ*., 82: 346-353.

Cusa, E., Obradors, N., Baldoma, L., Badia, J., Aguilar, J., 1999. Genetic analysis of a chromosomal region containing genes required for assimilation of allantoin nitrogen and linked glyoxylate metabolism in Escherichia coli. *J. Bacteriol*., 181: 7479-7484.

Davies, J.E., 1997. Origins, acquisition and dissemination of antibiotic resistance determinants. *Ciba. Found. Symp*., 207: 15-35.

De Lappe, N., Doran, G., O'Connor, J., O'Hare, C., Cormican, M., 2009. Characterization of bacteriphages used in the Salmonella enterica serovar Enteritidis phage-typing scheme. *J. Med. Microbiol.*, 58: 86-93.

Doublet, B., Boyd, D., Mulvey, M.R., Cloeckaert, A., 2005a. The Salmonella genomic island 1 is an integrative mobilizable element. *Mol. Microbiol.*, 55: 1911-1924.

Doublet, B., Carattoli, A., Whichard, J.M., White, D.G., Baucheron, S., Chaslus-Dancla, E., Cloeckaert, A., 2004. Plasmid-mediated florfenicol and ceftriaxone resistance encoded by the floR and bla(CMY-2) genes in Salmonella enterica serovars Typhimurium and Newport isolated in the United States. *FEMS. Microbiol. Lett.*, 233: 301-305.

Doublet, B., Golding, G.R., Mulvey, M.R., Cloeckaert, A., 2008. Secondary chromosomal attachment site and tandem integration of the mobilizable Salmonella genomic island 1. *PLoS ONE*., 3(4): e2060.

Doublet, B., Schwarz, S., Kehrenberg, C., Cloeckaert, A., 2005b. Florfenicol resistance gene floR is part of a novel transposon. *Antimicrob. Agents. Chemother*., 49: 2106-2108.

Dougan, G., Barrow, P., Achtman, M. S. enterica Typhimurium DT104 NCTC 13348

Dougan, G., Barrow, P., Achtman, M. *S. enterica Typhimurium* SL1344 NCTC 13347

Dougan, G., Barrow, P., Achtman, M. *Salmonella bongori* 12419 ATCC 43975

Economic Research Service, 2008. *Foodborne Illness Cost Calculator*. United States Department of Agriculture. Available from http://www.ers.usda.gov/publications.aspx [Accessed 12 March 2010].

Farr, S.B., Kogoma, T., 1991, Oxidative Stress Responses in Escherichia coli and Salmonella Typhimurium. *Microb. Rev.*, 55: 561-585.

Faucher, S.P., Porwollik, S., Dozois, C. M., McClelland, M., Daigle, F., 2005. Transcriptome of Salmonella enterica serovar Typhi within macrophages revealed through the selective capture of transcribed sequences. *Proc. Natl. Acad. Sci.*, 103: 1906-1911.

Felix, A., Callow, B.R., 1943. Typing of the paratyphoid B bacilli by means of Vi-bacteriophage. *Brit. Med. J*., 2: 127.

Felix, A., Pitt, R.M., 1934. A new antigen of B. typhosus. Its relation to virulence and to active and passive immunization. *Lancet*., 112: 186-191.

Fields, P.I., Swanson, R.V., Haidaris, C.G., Heffron, F., 1986. Mutants of Salmonella Typhimurium that cannot survive within the macrophage are avirulent. *Proc. Natl. Acad. Sci. USA.*, 83: 5189-5193.

Figueroa-Bossi, N., Bossi, L., 1999. Inducible prophages contribute to Salmonella virulence in mice. *Mol. Microbiol*., 33: 167-176.

Fisher, J.R., Eakin, R.E., 1957. Nitrogen excretion in developing chick embryos. *J. Embryol. Exp.Morphol*., 5: 215-224.

Fluit, A.C., Schmitz, F.J., 1999. Class 1 integrons, gene cassettes, mobility, and epidemiology. *Eur. J. Clin. Microbiol. Infect. Dis*., 18: 761-770.

Foley, S.L., Lynne, A. M., Nayak, R., 2008. Salmonella challenges: Prevalence in swine and poultry and potential pathogenicity of such isolates. *J. Anim. Sci*., 86: E149-E162.

Folkesson, A., Lofdahl, S., Normark, S., 2002. The Salmonella enterica subspecies I specific centisome 7 genomic island encodes novel protein families present in bacteria living in close contact with eukaryotic cells. *Res. Microbiol*., 153: 537-545.

Foster, J.W., Hall, H.K., 1991. Inducible pH homeostasis and the acid tolerance response of Salmonella Typhimurium. *J. Bacteriol*., 173: 5129-5135.

Fujiwara, S., Noguchi, T., 1995. Degradation of purines: only ureidoglycollate lyase out of four allantoin-degrading enzymes is present in mammals. *Biochem. J*., 312 (1): 315-318.

Garcia-del Portillo, F., Foster, J.W., Maguire, M.E., Finlay, B.B., 1992. Characterization of the micro-environment of Salmonella Typhimurium containing vacuoles within MDCK epithelial cells. *Mol. Microbiol*., 6: 3289-3297.

Garcia-Quintanilla, M., Casadesus, J., 2011. Virulence plasmid interchange between strains ATCC 14028, LT2, and SL1344 of Salmonella enterica serovar Typhimurium. *Plasmid*., 65: 169-175.

Gauss, P., Park, K., Spencer, T.E., Hacker, K.J., 1994. DNA helicase requirements for DNA replication during bacteriophage T4 infection. *J. Bacteriol.*, 176: 1667-1672.

Gerlach, R.G., Jäckel, D., Geymeier, N., Hensel, M., 2007. Salmonella pathogenicity island 4-mediated adhesion is coregulated with invasion genes in Salmonella enterica. *Infect. Immun*., 75: 4697-4709.

Giannella, R.A., 1979. Importance of the intestinal inflammatory reaction in Salmonella-mediated intestinal secretion. *Infect. Immun*., 23: 140-145.

Giraud, E., Cloeckaert, A., Kerboeuf, D., Chaslus-Dancla, E., 2000. Evidence for active efflux as the primary mechanism of resistance to ciprofloxacin in Salmonella enterica serovar Typhimurium. *Antimicrob. Agents. Chemother*., 44: 1223-1228.

Goodell, E.W., Higgins, C. F., 1987. Uptake of cell wall peptides by Salmonella Typhimurium and Escherichia coli. *J. Bacteriol*., 169: 3861-3865.

Gordon, M.A., 2008. Salmonella infections in immunocompromised adults. *J. Infect*., 56: 413-422.

Groisman, E.A., 2001. The pleiotropic two-component regulatory system PhoP-PhoQ. *J. Bacteriol*., 183: 1835-1842.

Groisman, E.A., Ochman, H., 1993. Cognate gene clusters govern invasion of host epithelial cells by Salmonella Typhimurium and Shigella flexneri. *EMBO. J.*, 12: 3779-3787.

Groisman, E.A., Ochman, H., 1997. How Salmonella became a pathogen. *Trends. Microbiol*., 5: 343-349.

Guerra, B., Soto, S., Helmuth, R., Mendoza, M.C., 2002. Characterization of a self-transferable plasmid from Salmonella enterica serotype Typhimurium clinical isolates carrying two integron-borne gene cassettes together with virulence and drug resistance genes. *Antimicrob. Agents. Chemother*., 46: 2977-2981.

Gupta, A., Fontana, J., Crowe, C., Bolstorff, B., Stout, A., Van Duyne, S., Hoekstra, M.P., Whicard, J.M., Barrett, T.J., Angulo, F.J., 2003. Emergence of multidrug-resistant Salmonella enterica serotype Newport infections resistant to expanded-spectrum cephalosporins in the United States. *J. Infect. Diseases*., 188: 1707-1716.

Gutnick, D., Calvo, J.M., Klopotowski, T., Ames, B.N., 1969. Compounds which serve as the sole source of carbon or nitrogen for Salmonella Typhimurium LT-2. *J. Bacteriol.,* 100: 215-219.

Heffernan, E.J., Reed, S., Hackett, J., Fierer, J., Roudier, C., Guiney, D., 1992. Mechanism of resistance to complement-mediated killing of bacteria encoded by the Salmonella Typhimurium virulence plasmid gene rck. *J. Clin. Invest.,* 90: 953-964.

Helms, M., Ethelberg, S., Mølbak, K., 2005. International Salmonella Typhimurium DT104 Infections, 1992-2001. *Emerg. Infect. Dis*., 11: 859-867.

Hensel, M., 2000. Salmonella pathogenicity island 2. *Mol. Microbiol*., 36: 1015-1023.

Hensel, M., 2004. Evolution of pathogenicity islands of Salmonella enterica. *Int. J. Med. Microbiol*., 294: 95-102.

Hermans, A.P.H.M., Abee, T., Zwietering, M.H., Aarts, H.J.M., 2005. Identification of novel Salmonella enterica serovar Typhimurium DT104-specific prophage and nonprophage chromosomal sequences among serovar Typhimurium isolates by genomic subtractive hybridisation. *App. Env. Microb.*, 71: 4979-4985.

Hiles, I.D., Higgins, C.F., 1986. Peptide uptake by Salmonella Typhimurium. *Eur. J. Biochem*., 158: 561-567.

Jana, S., Deb, J.K., 2005. Kinetic mechanism of streptomycin adenylyltransferase from a recombinant Escherichia coli. *Biotech. Lett*., 27: 519-524.

Joiner, K.A., 1988. Complement evasion by bacteria and parasites. *Ann. Rev. Microbiol.,* 42: 201-230.

Jones, B.D., Ghori, N., Falkow, S., 1994. Salmonella Typhimurium initiates murine infection by penetrating and destroying the specialized epithelial M cells of the Peyer's patches. *J. Exp. Med*., 180: 15-23.

Jones, B.D., Paterson, H.F., Hall, A., Falkow, S., 1993. Salmonella Typhimurium induces membrane ruffling by a growth factor-receptor-independent mechanism. *Proc. Natl. Acad. Sci. USA*., 90: 10390-10394.

Jones, W.R., Barcak, G.J., Wolf, R.E. Jr., 1990. Altered growth-rate-dependent regulation of 6-phosphogluconate dehydrogenase level in hisT mutants of Salmonella Typhimurium and Escherichia coli. *J. Bacteriol*., 172: 1197-1205.

Jones, Y.E., Mclaren, I. M., Wray, C., 2000. Laboratory aspects of Salmonella. In *Salmonella in Domestic Animals*. Wray, C., Wray, A. Eds., CABI Publishing, pp. 393-405.

Joyce, E.A., Chan, K., Salama, N.R., Falkow, S., 2002. Redefining bacterial populations: a post-genomic reformation. *Nat. Rev. Genetics*., 3: 462-473.

Kucken, D., Feucht, H., Kaulfers, P., 2000. Association of qacE and qacEDelta1 with multiple resistance to antibiotics and antiseptics in clinical isolates of Gram-negative bacteria. *FEMS. Microbiol. Lett.,* 183: 95-98.

Lacroix, F.J., Cloeckaert, A., Grepinet, O., Pinault, C., Popoff, M.Y., Waxin, H., Pardon, P., 1996. Salmonella Typhimurium acrB-like gene: identification and role in resistance to biliary salts and detergents and in murine infection. *FEMS. Microbiol. Lett.,* 135: 161-167.

Lan, R., Reeves, P.R., Octavia, S., 2009. Population structure, origins and evolution of major Salmonella enterica clones. *Infect. Genet. Evol*., 9: 996-1005.

Lawson, A.J., Desai, M., O'Brien, S.J., Davies, R.H., Ward, L.R., Threlfall, E.J., 2004. Molecular characterisation of an outbreak strain of multiresistant Salmonella enterica serovars Typhimurium DT104 in the UK. *Clin. Microbiol. Infect*., 10: 143-147.

Levings, R.S., Djordhevic, S.P., Hall, R. M., 2008. SGI2, a relative of Salmonella genomic island SGI1 with an independent origin. *Antimicrob. Agents. Chemother*., 52: 2529-2537.

Levings, R.S., Partridge, S.R., Djordjevic, S.P., Hall, R.M., 2007. SGI1-K, a variant of the SGI1 genomic island carrying a mercury resistance region, in Salmonella enterica serovar Kentucky. *Antimicrob. Agents. Chemother*., 51: 317-323.

Liebana, E., Garcia-Migura, L., Clouting, C., Clifton-Hadley, F.A., Lindsay, E., Threlfall, E.J., McDowell, S.W., Davies, R.H., 2002. Multiple genetic typing of Salmonella enterica serotype Typhimurium isolates of different phage types (DT104, U302, DT204b, and DT49) from animals and humans in England, Wales, and Northern Ireland. *J. Clin. Microbiol*., 40: 4450-4456.

Lopilato, J.E., Garwin, J.L., Emr, S.D., Silhavy, T.J., Beckwith, J.R., 1984. D-ribose metabolism in Escherichia coli K-12: genetics, regulation, and transport. *J. Bacteriol*., 158: 665-673.

Lorenz, M.C., Fink, G. R., 2002. Life and death in a macrophage: role of the glyoxylate cycle in virulence. *Eukaryotic. Cell*., 1: 657-662.

Lucarelli, C., Dionisi, A.M., Torpdahl, M., Villa, L., Graziani, C., Hopkins, K., Threlfall, J., Caprioli, A., Luzzi, I., 2010. Evidence for a second genomic island conferring multidrug resistance in a clonal group of strains of Salmonella enterica serovar Typhimurium and its monophasic variant circulating in Italy, Denmark, and the United Kingdom. *J. Clin. Microbiol*., 48: 2103-2109.

Lucas, R.L., Lee, C.A., 2001. Roles of hilC and hilD in regulation of hilA expression in Salmonella enterica serovar Typhimurium. *J. Bacteriol*., 183: 2733-2745.

Ma, D., Cook, D.N., Alberti, M., Pon, N.G., Nikaido, H., Hearst, J.E., 1995. Genes acrA and acrB encode a stress-induced efflux system of Escherichia coli. *Mol. Microbiol*., 16: 45-55.

Madigan, M.T., Martinko, J.M., Brock, T.D.B., 2006. *Brock biology of microorganisms*. 11th ed., Pearson Prentice Hall: Upper Saddle River, NJ.

Maloy, S.R., Bohlander, M., Nunn, W.D., 1980. Elevated levels of glyoxylate shunt enzymes in Escherichia coli strains constitutive for fatty acid degradation. *J. Bacteriol*., 143: 720-725.

Matiasovicova, J., Adams, P., Barrow, P.A., Hradecka, H., Malcova, M., Karpiskova, R., Budinska, E., Pilousova, L., Rychlik, I., 2007. Identification of putative ancestors of the multidrug-resistant Salmonella enterica serovar Typhimurium DT104 clone harboring the Salmonella genomic island 1. *Arch. Microbiol*., 187: 415-424.

Matthew, M., Sykes, R.B., 1977. Properties of the beta-lactamase specified by the Pseudomonas plasmid RPL11. *J. Bacteriol*., 132: 341-345.

Maurelli, A.T., Fernandez, R.E., Bloch, C.A., Rode, C.K., Fasano, A., 1998. Black holes and bacterial pathogenicity: a large genomic deletion that enhances the virulence of Shigella spp. and enteroinvasive Escherichia coli. *Proc. Natl. Acad. Sci. USA*., 95: 3943-3948.

McClelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W., Stoneking, T., Nhan, M., Waterston, R., Wilson, R.K., 2001. Complete genome sequence of Salmonella enterica serovar Typhimurium LT2. *Nature*., 413: 852-856.

McMurry, L., Petrucci, R.E. Jr., Levy, S.B., 1980. Active efflux of tetracycline encoded by four genetically different tetracycline resistance determinants in Escherichia coli. *Proc. Natl. Acad. Sc.i USA*., 77: 3974-3977.

McNally, A., Dalton, T., La Ragione, R.M., Stapleton, K., Manning, G., Newell, D.G., 2006. Yersinia enterocolitica isolates of differing biotypes from humans and

animals are adherent, invasive and persist in macrophages, but differ in cytokine secretion profiles in vitro. *J. Med. Microb*., 55: 1725-1734.

McNally, A., La Ragione, R.M., Best, A., Manning, G., Newell, D.G., 2007. An aflagellate mutant Yersinia enterocolitica biotype 1A strain displays altered invasion of epithelial cells, persistence in macrophages, and cytokine secretion profiles in vitro. *Microbiol*., 153: 1339-1349.

McQuiston, J.R., Herrera-Leon, S., Wertheim, B.C., Doyle, J., Fields, P.I., Tauxe, R.V., Logsdon, J.M., 2008. Molecular phylogeny of the salmonellae: relationships among Salmonella species and subspecies determined from four housekeeping genes and evidence of lateral gene transfer events. *J. Bacteriol*., 190: 7060-7067.

Mead, P.S., Slutsker L., Dietz, V., McCraig, L.F., Bresee, J.S., Shapiro, C., Griffin, P.M., Tauxe, R.V., 1999. Food related illness and death in the United States. *Emerg. Infect. Dis*., 5: 607-625.

Mills, D.M., Bajaj, V., Lee, C.A., 1995. A 40 kb chromosomal fragment encoding Salmonella Typhimurium invasion genes is absent from the corresponding region of the Escherichia coli K-12 chromosome. *Mol. Microbiol*., 15: 749-759.

Morgan, E., 2007. Salmonella Pathogenicity Islands. In *Salmonella: Molecular Biology and Pathogenesis*. Rhen, M., Maskell, D., Mastroeni, P., Threlfall, J. eds., Horizon Bioscience, pp. 67-88.

Mulvey, M.R., Boyd, D.A., Olson, A.B., Doublet, B., Cloeckaert, A., 2006. The genetics of Salmonella genomic island 1. *Microbes. Infect*., 8: 1915-1922.

Neidhardt, F.C., Bloch, P.L., Smith, D.F., 1974. Culture medium for enterobacteria. *J. Bacteriol*., 119: 736-747.

Ochman, H., Groisman, E.A., 1996. Distribution of pathogenicity islands in Salmonella spp. *Infect. Immun*., 64: 5410-5412.

Ornston, L.N., Ornston, M.K., 1969. Regulation of glyoxylate metabolism in Escherichia coli K-12. *J. Bacteriol*., 98: 1098-1108.

Pan, Z., Carter, B., Nunez-Garcia, J., Abuoun, M., Fookes, M., Ivens, A., Woodward, M.J., Anjum, M.F., 2009. Identification of genetic and phenotypic differences associated with prevalent and non-prevalent Salmonella Enteritidis phage types: analysis of variation in amino acid transport. *Microbiology.*, 155: 3200-3213.

Parkhill, J., Dougan, G., James, K.D., Thomson, N.R., Pickard, D., Wain, J., Churcher, C., Mungall, K.L., Bentley, S.D., Holden, M.T., Sebaihia, M., Baker, S., Basham, D., Brooks, K., Chillingworth, T., Connerton, P., Cronin, A., Davis, P., Davies, R.M., Dowd, L., White, N., Farrar, J., Feltwell, T., Hamlin, N., Haque, A., Hien, T.T., Holroyd, S., Jagels, K., Krogh, A., Larsen, T.S., Leather, S., Moule, S., O'Gaora, P., Parry, C., Quail, M., Rutherford, K., Simmonds, M., Skelton, J., Stevens, K., Whitehead, S., Barrell, B.G., 2001. Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18. *Nature*., 413: 848-852.

Pascopella, L., Raupach, B., Ghori, N., Monack, D., Falkow, S., Small, P.L.C., 1995. Host restriction phenotypes of Salmonella typhi and Salmonella Gallinarum. *Infect. Immun*., 63: 4329-4335.

Phan, M.D., Kidgell, C., Nair, S., Holt, K.E., Turner, A.K., Hinds, J., Butcher, P., Cooke, F.J., Thomson, N.R., Titball, R., Bhutta, Z.A., Hasan, R., Dougan, G., Wain, J., 2009. Variation in Salmonella enterica serovar typhi IncHI1 plasmids during the global spread of resistant typhoid fever. *Antimicrob. Agents. Chemother*., 53: 716-727.

Pickard, D., Wain, J., Baker, S., Line, A., Chohan, S., Fookes, M., Barron, A., Gaora, P.O., Chabalgoity, J.A., Thanky, N., Scholes, C., Thomson, N., Quail, M., Parkhill, J., Dougan, G., 2003. Composition, acquisition, and distribution of the Vi exopolysaccharide-encoding Salmonella enterica pathogenicity island SPI-7. *J. Bacteriol*., 185: 5055-5065.

Plym Forshell, L., Wierup, M., 2006. Salmonella contamination: a significant challenge to the global marketing of animal food products. *Rev. Sci. Tech. Off. Int. Epiz*., 25: 541-554.

Popoff, M.Y., Bockemühl, J., Gheesling, L.L., 2001. Antigenic formulas of the Salmonella serovars. In *World Health Organisation Collaborating Centre for Reference and Research on Salmonella*, Pasteur Institute.

Popoff, M.Y., Bockemühl, J., Gheesling, L.L., 2002. Supplement 2002 (no. 46) to the Kauffmann-White scheme. *Res. Microbiol*., 155: 568-570.

Porwollik, S., Boyd, E.F., Choy, C., Cheng, P., Florea, L., Proctor, E., McClelland, M., 2004. Characterization of Salmonella enterica subspecies I genovars by use of microarrays. *J. Bacteriol*., 186: 5883-5898.

Powling, J., 2005. *National Enteric Pathogens Surveillance Scheme. Non-Human Annual Report 2004*. The University of Melbourne, Melbourne, Australia.

Price-Carter, M., Fazzio, T.G., Vallbona, E.I., Roth, J.R., 2005. Polyphosphate kinase protects Salmonella enterica from weak organic acid stress. *J. Bacteriol*., 187: 3088-3099.

Reeves, M.W., Evins, G.M., Heiba, A.A., Plikaytis, B.D., Farmer III, J.J., 1989. Clonal nature of Salmonella typhi and its genetic relatedness to other salmonellae as shown by multilocus enzyme electrophoresis, and proposal of Salmonella bongori comb. nov. *J. Clin. Microbiol*., 27: 313-320.

Refsum, T., Handeland, K., Baggesen, D.L., Holstad, G., Kapperud, G., 2002. Salmonellae in avian wildlife in Norway from 1969 to 2000. *App. Env. Microb*., 68: 5595-5599.

Rest, R.F., Cooney, M. H., Spitznagel, J.K., 1977. Susceptibility of lipopolysaccharide mutants to the bactericidal action of human neutrophil lysosomal fractions. *Infect. Immun*., 16: 145-151.

Rhen, M., Sukupolvi, S., 1988. The role of the traT gene of the Salmonella Typhimurium virulence plasmid for serum resistance and growth within liver macrophages. *Microbial. Path*., 5: 275-285.

Rodrigue, D.C., Tauxe, R.V., Rowe, B., 1990. International increase in Salmonella enteritidis: A new pandemic? *Epid. Infect*., 105: 21-27.

Roof, D.M., Roth, J.R., 1989. Functions required for vitamin B12-dependent ethanolamine utilization in Salmonella Typhimurium. *J. Bacteriol*., 171: 3316-3323.

Roth, J.R., Lawrence, J.G., Rubenfield, M., Kieffer-Higgins, S., Church, G.M., 1993. Characterization of the cobalamin (vitamin B12) biosynthetic genes of Salmonella Typhimurium. *J. Bacteriol*., 175: 3303-3316.

Saitoh, M., Tanaka, K., Nishimori, K., Makino, S., Kanno, T., Ishihara, R., Hatama, S., Kitano, R., Kishima, M., Sameshima, T., Akiba, M., Nakazawa, M., Uchida, I., 2005. The *artAB* genes encode a putative ADP-ribosyltransferase toxin homologue associated with Salmonella enterica serovar Typhimurium DT104. *Microbiology*., 151: 3089-3096.

Sandvang, D., Møller Aarestrup, F., Bogø Jensen, L., 1998. Characterisation of integrons and antibiotic resistance genes in Danish multiresitant Salmonella enterica Typhimurium DT104. *FEMS. Micro. Lett*., 160: 37-41.

Santiviago, C.A., Blondel, C.J., Quezada, C.P., Silva, C.A., Tobar, P.M., Porwollik, S., McClelland, M., Andrews-Polymenis, H.L., Toro, C.S., Zaldivar, M., Contreras, I., 2010. Spontaneous excision of the Salmonella enterica serovar Enteritidis-specific defective prophage-like element phiSE14. *J. Bacteriol*., 192: 2246-2254.

Saunders, M.P., Wu, G., Abuoun, M., Pan, Z., Anjum, M., Woodward, M.J., 2011. Optical genetic mapping defines regions of chromosomal variation in serovars of S. enterica subsp. enterica of concern for human and animal health. *Epidemiol. Infect.*, 7: 1065-1074.

Schechter, L.M., Damrauer, S.M., Lee, C.A., 1999. Two AraC/XylS family members can independently counteract the effect of repressing sequences upstream of the hilA promoter. *Mol. Microbiol*., 32: 629-642.

Schnabel, E.L., Jones, A.L., 1999. Distribution of tetracycline resistance genes and transposons among phylloplane bacteria in Michigan apple orchards. *Appl. Environ. Microbiol*., 65: 4898-4907.

Schwan, W.R., Huang, X.Z., Hu, L., Kopecko, D.J., 2000. Differential bacterial survival, replication, and apoptosis-inducing ability of Salmonella serovars within human and murine macrophages. *Infect. Immun*., 68: 1005-1013.

Shaibe, E., Metzer, E., Halpern, Y.S., 1985a. Control of utilization of L-arginine, L-ornithine, agmatine, and putrescine as nitrogen sources in Escherichia coli K-12. *J. Bacteriol*., 163: 938-942.

Shaibe, E., Metzer, E., Halpern, Y.S., 1985b. Metabolic pathway for the utilization of L-arginine, L-ornithine, agmatine, and putrescine as nitrogen sources in Escherichia coli K-12. *J. Bacteriol*., 163: 933-937.

Sharp, M.W., Rawson, B.C., 1992. Persistent Salmonella Typhimurium PT104 infection in a dairy herd. *Vet. Rec*., 131: 375-376.

Skold, O., 2001. Resistance to trimethoprim and sulfonamides. *Vet. Res*., 32: 261-273.

Soong, G., Muir, A., Gomez, M.I., Waks, J., Reddy, B., Planet, P., Singh, P.K., Kaneko, Y., Wolfgang, M.C., Hsiao, Y.S., Tong, L., Prince, A., 2006. Bacterial neuraminidase facilitates mucosal infection by participating in biofilm production. *J. Clin. Invest*., 116: 2297-2305.

Soyer, Y., Moreno Switt, A., Davis, M.A., Maurer, J., McDonough, P.L., Schoonmaker-Bopp, D.J., Dumas, N.B., Root, T., Warnick, L.D., Grohn, Y.T., Wiedmann, M., 2009. Salmonella enterica serotype 4,5,12:i:-, an emerging Salmonella serotype that represents multiple distinct clones. *J. Clin. Microbiol*., 47: 3546-3556.

Tanaka, K., Nishimori, K., Makino, S., Nishimori, T., Kanno, T., Ishihara, R., Sameshima, T., Akiba, M., Nakazawa, M., Yokomizo, Y., Uchida, I., 2004. Molecular characterization of a prophage of Salmonella enterica serotype Typhimurium DT104. *J. Clin. Microbiol*., 42: 1807-1812.

Thomson, N., Baker, S., Pickard, D., Fookes, M., Anjum, M., Hamlin, N., Wain, J., House, D., Bhutta, Z., Chan, K., Falkow, S., Parkhill, J., Woodward, M., Ivens, A., Dougan, G., 2004. The role of prophage-like elements in the diversity of Salmonella enterica serovars. *J. Mol. Biol*., 339: 279-300.

Thomson, N.R., Clayton, D.J., Windhorst, D., Vernikos, G., Davidson, S., Churcher, C., Quail, M.A., Stevens, M., Jones, M.A., Watson, M., Barron, A., Layton, A., Pickard, D., Kingsley, R.A., Bignell, A., Clark, L., Harris, B., Ormond, D., Abdellah, Z., Brooks, K., Cherevach, I., Chillingworth, T., Woodward, J., Norberczak, H., Lord, A., Arrowsmith, C., Jagels, K., Moule, S., Mungall, K., Sanders, M., Whitehead, S., Chabalgoity, J.A., Maskell, D., Humphrey, T., Roberts, M., Barrow, P.A., Dougan, G., Parkhill, J., 2008. Comparative genome analysis of Salmonella Enteritidis PT4 and Salmonella Gallinarum 287/91 provides insights into evolutionary and host adaptation pathways. *Genome. Res*., 18: 1624-1637.

Threlfall, E.J., Hampton, M.D., Schofield, S.L., 1997. Investigating Salmonella Typhimurium DT104 infections. *Comm. Dis. Rep*., 7: R137 R140.

Threlfall, E.J., Frost, J.A., Ward, L.R., Rowe, B., 1994. Epidemic in cattle and humans of Salmonella Typhimurium DT104 with chromosomally integrated multiple drug resistance. *Vet. Rec*., 134: 577.

Threlfall, E.J., Hall, M.L.M., Rowe, B., 1992. Salmonella bacteraemia in England and Wales. *J. Clin. Pathol*., 45: 34-36.

Threlfall, E.J., Hampton, M. D., Chart, H., Hopkins, K., Ward, L.R., Tebbutt, G., 2004. Emergence of new subclones of multiresistant S. Typhimurium DT104 possibly associated with poultry meat. *Vet. Rec.*, 154: 89-90.

Threlfall, E.J., Ward, L.R., Rowe, B., 1997. Increasing incidence of resistance to trimethoprim and ciprofloxacin in epidemic Salmonella Typhimurium DT104 in England and Wales. *Euro. Surv*., 2: 81-84.

Turner, A.K., Lovell, M.A., Hulme, S.D., Zhang-Barber, L., Barrow, P.A., 1998. Identification of Salmonella Typhimurium genes required for colonization of the chicken alimentary tract and for virulence in newly hatched chicks. *Infect. Immun*., 66: 2099-2106.

Uchida, I., Ishihara, R., Tanaka, K., Hata, E., Makino, S., Kanno, T., Hatama, S., Kishima, M., Akiba, M., Watanabe, A., Kubota, T., 2009. Salmonella enterica serotype Typhimurium DT104 ArtA-dependent modification of pertussis toxin-sensitive G proteins in the presence of [^{32}P]NAD. *Microbiology*., 155: 3710-3718.

Velge, P., Cloeckaert, A., Barrow, P., 2005. Emergence of Salmonella epidemics: the problems related to Salmonella enterica serotype Enteritidis and multiple antibiotic resistance in other major serotypes. *Vet. Res*., 36: 267-288.

VLA, 2007. *Salmonella in Livestock Production in GB*. Veterinary Laboratories Agency, Department of Environment, Food and Rural Affairs, United Kingdom.

Voetsch, A.C., Van Gilder, T.J., Angulo, F.J., Farley, M. M., Shallow, S., Marcus, R., Cieslak, P.R., Deneen, V.C., Tauxe, R.V., 2004. FoodNet estimate of the burden of illness caused by nontyphoidal Salmonella infections in the United States. *Clin. Infect. Dis.*, S127-134.

Wain, J., Diem Nga, L.T., Kidgell, C., James, K., Fortune, S., Song Diep, T., Ali, T., Peadar, O.G., Parry, C., Parkhill, J., Farrar, J., White, N.J., Dougan, G., 2003. Molecular analysis of incHI1 antimicrobial resistance plasmids from Salmonella serovar Typhi strains associated with typhoid fever. *Antimicrob. Agents. Chemother*., 47: 2732-2739.

Wall, P.G., Morgan, D., Lamden, K., Ryan, M., Griffin, M., Threlfall, E.J., Ward, L.R., Rowe, B., 1994. A case control study of infection with an epidemic strain of multi-resistant Salmonella Typhimurium DT104 in England and Wales. *Comm. Dis. Rep*., 4: 130-135.

Webb, V., Davies, J., 1994. Accidental release of antibiotic-resistance genes. *Trends. Biotechnol*., 12: 74-75.

Weissbach, A., Sprinson, D.B., 1953a. The metabolism of 2-carbon compounds related to glycine. I. Glyoxylic acid. *J. Biol. Chem*., 203: 1023-1030.

Weissbach, A., Sprinson, D.B., 1953b. The metabolism of 2-carbon compounds related to glycine. II. Ethanolamine. *J. Biol. Chem*., 203: 1031-1037.

WHO, 2003. *WHO Surveillance Programme for Control of Foodborne Infections and Intoxications in Europe 8th Report 1999-2000*. Bfr - FAO/WHO Collaborating Centre for Research and training in Food Hygiene and Zoonoses, Berlin, Germany.

Wilson, R.L., Elthon, J., Clegg, S., Jones, B.D., 2000. Salmonella enterica serovars Gallinarum and pullorum expressing Salmonella enterica serovar Typhimurium type 1 fimbriae exhibit increased invasiveness for mammalian cells. *Infect. Immun*., 68: 4782-4785.

Winkler, M.E., Roth, D.J., Hartman, P.E., 1978. Promoter- and attenuator-related metabolic regulation of the Salmonella Typhimurium histidine operon. *J. Bacteriol*. 133: 830-843.

Woodward, M.J., Charles, H.P., 1982. Genes for l-sorbose utilization in Escherichia coli. *J. Gen. Microbiol*., 128: 1969-1980.

Wray, C., McLaren, I.M., Jones, Y.E., 1998. The epidemiology of Salmonella Typhimurium in cattle: plasmid profile analysis of definitive phage type (DT) 204c. *J. Med. Microbiol*., 47: 483-487.

Wray, C., McLaren, I., Parkinson, N.M., Beedell, Y., 1987. Differentiation of Salmonella Typhimurium DT204c by plasmid profile and biotyping. *Vet. Rec*., 121: 514-516.

Wray, C., Sojka, W.J., 1978. Experimental Salmonella Typhimurium infection in calves. *Res. Vet. Sci*., 25: 139-143.

Wright, J.M., Boyle, S.M., 1984. Intergeneric homology of the speC gene encoding biosynthetic ornithine decarboxylase in Escherichia coli. *J. Bacteriol*., 159: 1074-1076.

Yoshida, H., Bogaki, M., Nakamura, M., Yamanaka, L.M., Nakamura, S., 1991. Quinolone resistance-determining region in the DNA gyrase gyrB gene of Escherichia coli. *Antimicrob. Agents. Chemother*., 35: 1647-1650.

Zhou, S., Kile, A., Bechner, M., Place, M., Kvikstad, E., Deng, W., Wei, J., Severin, J., Runnheim, R., Churas, C., Forrest, D., Dimalanta, E.T., Lamers, C., Burland, V., Blattner, F. R., Schwartz, D.C., 2004. Single-molecule approach to bacterial genomic comparisons via optical mapping. *J. Bacteriol*., 186: 7773-7782.